DOCUMENT RESUME

ED 217 079 TM 820 337

AUTHOR Brzezinski, Evelyn; Demaline, Randy
TITLE An Empirical Investigation of the Results of Two Different Test Development Strategies. Final Report for the Period March 15, 1981 - December 31, 1981.
INSTITUTION Northwest Regional Educational Lab., Portland, Oreg.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Jan 82
NOTE 220p.; Some sections may be marginally legible due to small print.

EDRS PRICE MF01/PC09 Plus Postage.
DESCRIPTORS *Comparative Analysis; Computation; Cost Effectiveness; *Criterion Referenced Tests; Difficulty Level; Grade 4; Intermediate Grades; *Item Banks; Reading Comprehension; *Test Construction; *Test Items
IDENTIFIERS *Test Specifications

ABSTRACT

Item banks offer a different approach to test development than that of the more traditional item specification method. This study was designed as an initial comparative investigation in the Area. Fourth grade math and reading tests were developed using both items written from item specifications and items drawn from an existing item bank. Four separate panels were formed to construct the tests and their development activities were documented for time and cost comparisons. Psychometric properties of the four tests were also compared although no consistent results between the two general methods were found. Finally, overall quality and content coverage of the tests were judged by classroom teachers. The report is divided into five major sections. The introduction includes a statement of the problem under investigation in the study, a review of related literature, and a statement of the current status of item banking in education. The second section, Design of the Study, describes the plans for the study. The third section, Conduct of the Study, explains how the study actually was conducted, showing when and how the actual implementation differed from the planned implementation. The fourth section presents Results of the Study. The last section discusses what has been learned from the research and talks about Implications for Future Test Development Efforts, which include (1) necessity of reading level indicator from reading passages; (2) the number of items available in a reading passage; and (3) information concerning the difficulty of items within the bank. A number of appendices include the materials developed as part of this study. (Author/GK)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED217079

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

An Empirical Investigation of the Results of Two Different Test Development Strategies

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

M. M. Rogers

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

FINAL REPORT

for the period
March 15, 1981 through December 31, 1981

submitted January 1982

Northwest Regional Educational Laboratory
300 S.W. Sixth Avenue
Portland, Oregon 97204
Telephone: (503) 248-6800

TM 820 337

Table of Contents

Page

Acknowledgments i

List of Authors of Report Sections iii

Abstract iv

I. Introduction 1
- A. Statement of the Problem 2
- B. Review of Related Literature 3
- C. Current Status of Item Banking 4

II. Design of the Study 8
- A. Selection of Accounting Categories 8
- B. Operation of Content Panels 8
- C. Plans for Test Development 9
- D. Test Review, Piloting and Analysis Plans 9

III. Conduct of the Study 11
- A. Selection of Content Panels 11
- B. Selection of Test Content 12
- C. Operation of Item Specification Panels 12
- D. Operation of Item Bank Panels 19
- E. Content Review of Tests 23
- F. Pilot Testing 24

IV. Results of the Study 25
- A. Results from the Pilot Test: Assessment of Psychometric Properties 25
- B. Results from the Content Review: Assessment of Content Validity 29
- C. Results from the Time and Cost Analyses 32

V. Implications for Future Test Development Efforts 35

Appendices

Appendix A—Description of NWREL Item Bank 39

Appendix B—Letter to Potential Content Panel Members 63

Appendix C—Letter to Persons Selected as Content Panelists 65

ACKNOWLEDGMENTS

The authors of this report wish to acknowledge the assistance provided by a number of persons without whose help the study reported herein could not have been completed. First, thanks are due to the 18 individuals who served as content panel members:

Martha Elsken
Lianne Forney
Jayne Freeman
Joyce Hindman
Charles Jenkins
Mary Jubitz
Carolyn Kilmer
Mary Leverington
Sharon McCullough
Joan Miller
Velma Miller
Sharon Norlen
Joan Okuhara
Linda Powers
Phyllis Rodinsky
John Smith
Eleanor Tanner
Pam Wickman

The panel members worked conscientiously on their assignments, and we appreciate their professional approach to the tasks assigned to them.

A second group of persons who should be acknowledged are the evaluation directors of the five school districts which participated in the study:

Marilyn Averill, Boulder Valley (CO) School District

Walter DeBoer and Jill Wiser, Grand Rapids (MI) School District

David Hazen, Ingham County (MI) Intermediate School

Wayne Neuburger, Beaverton (OR) School District

Sandy Wilson, Central Valley (WA) School District

The assistance of these individuals in soliciting teachers to serve as test reviewers (in the case of Boulder, Grand Rapids, Ingham County and Beaverton) and to administer the tests to students (in Central Valley) was invaluable. We know that local school district staff are always being asked to participate in studies--many of which are of questionable utility to them. We hope the above individuals find this report useful, and we thank them for helping us in the study.

Several professional colleagues also provided valuable input. Sue Hiscox of the Northwest Regional Educational Laboratory provided some much needed help in writing the first draft of reading passages and questions for the Reading Item Specification Panel. Michael Hiscox of Interwest Applied Research assisted in various portions of the study, including consulting on design and procedures, contributing to the final report in the "Status of Item Banking" section and reviewing the entire report and offering suggestions for changes to make the results clearer. Thel Kocher, NIE project officer, guided the project during design and implementation, raised technical questions and provided comments that helped the authors clarify the draft report. Their help is gratefully acknowledged.

Finally, thanks are due to the support staff at the Northwest Regional Educational Laboratory without whose help this study would not have been completed. Georgette Schulberg, Debbie Van Horn, Pandora Pierce and Bonnie Holt all contributed with their time and expertise, and we appreciate their assistance.

List of Authors of Report Sections

	Section	Author
	Abstract	Randy Demaline
I.	Introduction	Evelyn Brzezinski
II.	Design of the Study	Evelyn Brzezinski
III.	Conduct of the Study	
	A. Selection of Contact Panels	Evelyn Brzezinski
	B. Selection of Test Content	Evelyn Brzezinski
	C. Operation of Item Specification Panels	Randy Demaline
	D. Operation of Item Bank Panels	Randy Demaline
	E. Content Review of Tests	Evelyn Brzezinski
	F. Pilot Testing	Evelyn Brzezinski
IV.	Results of the Study	Evelyn Brzezinski
V.	Implications for Future Test Development Efforts	Randy Demaline

ABSTRACT

Item banks offer a different approach to test development than that of the more traditional item specification method. Given the number of item banks that are appearing, test developers will begin to question the differences between the two techniques. This study was designed as an initial comparative investigation in the area. Fourth grade math and reading tests were developed using both items written from item specifications and items drawn from an existing item bank. Four separate panels were formed to construct the tests and their development activities were documented for time and cost comparisons. Psychometric properties of the four tests were also compared although no consistent results between the two general methods were found. Finally, overall quality and content coverage of the tests were judged by classroom teachers. Development stages noted during the item bank procedure highlighted the lack of standard procedures to use during item bank test development. Finally, this study presents implications for future test development efforts using item banks: (1) necessity of reading level indicators from reading passages; (2) the number of items available in a reading passage; and (3) information concerning the difficulty of items within the bank.

I. INTRODUCTION

Item banking is currently in transition from a technical specialty for a few measurement experts to a pragmatic necessity for nearly all test developers. Test developers, like all other members of the educational community, are being faced with massive reductions in available funding. They are looking for ways to get the most pay-off from their declining test development dollars. The appeal of item banks, where items already exist and a test developer simply has to select appropriate items for a specific test, is great. But the question of quality remains: Is a test developed from an item bank as "good" as a test developed using traditional methods such as writing objectives and item specifications, developing sample items, and so on? And is a test developed from an item bank really less expensive than one developed by traditional methods? The study described in this report was designed to answer the question of quality, as well as some of the questions related to the cost effectiveness of developing tests from item banks.

The report is divided into five major sections. This Introduction includes a statement of the problem under investigation in the study, a review of related literature, and a statement of the current status of item banking in education. The second section, Design of the Study, describes the plans for the study. The third section, Conduct of the Study, explains how the study actually was conducted, showing when and how the actual implementation differed from the planned implementation. The fourth section presents Results of the Study. The last section discusses what has been learned from the research and talks about Implications for Future Test Development Efforts. A number of appendices include the materials developed as part of this study.

This study is one piece of a larger research effort that should be undertaken to identify cost-effective methods of test development. It seems clear that the financial outlook for most of the nation's schools is going to get worse before it gets better. Assessment specialists in the schools are going to have to be able to support their requests for funding with data that show they are using those resources in the most efficient way possible. Not only must they show that they are spending school district dollars wisely, but that those dollars are buying valued services. The study described in this report does not address the question of value, but other researchers must address it soon. Rather, this study contributes to the body of knowledge about the effective use of resources. Without that information, questions of value seem premature. The results of this study will, hopefully, prompt others to consider the costs and benefits of the assessment and evaluation activities going on in school districts around the country.

Statement of the Problem

The study of the effectiveness of item banking is an idea whose time has come. In a study conducted by the Northwest Regional Educational Laboratory (NWREL) in Spring 1980, over 70 different K-12 educational test item banks were discovered across the country. Close to half of the state departments of education already have or are contemplating developing or acquiring item banks. Numerous local districts and intermediate school districts also have item banks. Although many of the existing item banks have not yet solved the problems of dissemination, at least the building blocks are present for streamlined test development activities. Such streamlining is critical in this time of increasingly tight funding cycles for education. Certainly item banks are viewed by many as a promising alternative to current test development practices.

Much literature exists on how to develop good objective referenced tests (ORTs), and there is considerable consistency in the recommended procedures. But that literature presumes that tests are being developed from scratch: content is being specified and items are being written to measure the desired content. Test development using item banks, on the other hand, is an area with very little research. Professional writings on the subject of item banks are divided into two major categories: either (1) how to organize a collection of items for item banking purposes or (2) how a particular item bank should be or has been used in practice. No one seems to be investigating how test development procedures should be adapted from current standard practice in order to take advantage of the fact that, with access to banks, items are already available and do not have to be written.

Item banks are viewed as a promising direction for test development because they appear to obviate the need for the very time consuming and expensive steps of developing item specifications (detailed requirements for prototype items) and then, from them, the items themselves. (Of course, those steps may have been followed when the items in the bank were first written.) But Popham, Hambleton and others tell us that those steps are critical to the development of high quality objective referenced *tests*, not just objective referenced test *items*.[1] According to those experts, item specifications are necessary to ensure high content validity while item development activities yield a feeling of ownership on the part of the persons who develop the tests. As long as people feel compelled to carry out both steps, the maximum utility of item banks cannot be realized.

---

[1] See, for example, chapters 5 and 6 in W. James Popham, *Criterion-referenced measurement*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1978. Also, R.K. Hambleton, *et al*., Criterion-referenced testing and measurement: A review of technical issues and developments. *Review of Educational Research*, 1978, *48*, 1-47.

What is needed is research to determine whether the item specification and item development steps are, in fact, necessary in order to produce ORTs with high content validity and psychometric quality. Clearly, this is a tremendous gap in the body of knowledge related to test development. If item banks are to achieve their full potential, it is a gap which must be closed by research studies such as the one described in this report.

Review of Related Literature

As previously noted, there is much literature available to guide persons developing objective referenced tests. Hambleton, in the presession on Criterion-Referenced Test Development and Validation Methods which he presented for the last several years prior to the annual meeting of the American Educational Research Association, emphasizes the need to prepare elaborate test item specifications to guide persons who will write test items for an objective (or criterion, in his terminology) referenced test. Hambleton draws heavily from the Popham book noted earlier when he lists the components of a good item specification, as follows:

"1. General Description - This provides an overview of the behavior that is to be tested and which will be described in detail later.

"2. Sample Test Item (including directions to the student about how to respond) - The sample item serves as an illustration for individuals unable to read the detailed description (because of time involved) and, more importantly, it provides format cues for the item writers. It specifies the preferred form in which the items can be constructed.

"3. Stimulus Attributes - The attributes of the stimulus materials are specified, along with delimitations on possible stimuli. Popham suggests three methods for specifying the content: (a) spell out rules or algorithms which are used in generating and delimiting the content; (b) list all the content that might be included; or (c) isolate and describe the defining attributes of all eligible content.

"4. Response Attributes - This section specifies whether the examinee is to select from a set of response options for a test question or to construct a response. This section should specify the rules/criteria upon which both sorts (if both are used) or a single sort of response type is to be treated.

"5. Specification Supplement - This section, if used, would contain information on the stimulus attributes and/or response attributes which, if included, would have made the respective sections too long. For instance, content listings might be included here." (Quoted as excerpts from Hambleton's presession materials.)[2]

Hambleton concludes this listing with the comment, "Following through these series of steps should help the test developer come up with a quite well-developed set of test specifications" (Hambleton presession materials, Unit 2, p. 16). That is no doubt true. But no one will argue the fact that it is a very time consuming process. Is item specification development really a necessary process when the items are already available in item banks?

An ERIC search conducted in Spring 1980 crossed ITEM BANKS with COMPUTERS and ITEM BANKS with TEST CONSTRUCTION. Citations for 126 articles and studies resulted. A review of those documents failed to produce any study which looked at the quality of tests produced from an item bank compared to the quality of tests developed from scratch. Thus, apparently people are using item banks to develop tests, but do not know (or at least do not report) how the quality of those tests compare with other tests developed in the way recommended by Popham, Hambleton and others.

Current Status of Item Banking

In the 1980 report A Guide to Item Banking in Education,[3] the point was made that there are hundreds of existing item collections and several hundred thousand existing test items. The report asked the question, "Why would anyone write another test item when so many already exist?" One reason might be that the psychometric qualities of item bank items may not be as good as those developed from item specifications--a possibility which the study under discussion attempts to resolve. But even if one ignores the potential measurement problems, item banking as a practice still does not demonstrate the vigor that might be expected from a logical analysis of the advantages that banks could provide.

---

[2] Hambleton, R.K., Eignor, D. and Swaminathan, H. Criterion-referenced test development and validation methods. AERA Presession materials, Unit 2, pp. 12-16. Presented in Toronto, March 26-27, 1978.

[3] Hiscox, M.D. and Brzezinski, E.J. A guide to item banking in education. Portland, OR: Northwest Regional Educational Laboratory, 1980.

When a lack of enthusiasm is suggested, it is not meant to imply any shortage of item banks or items. There are enough, and possibly too many, item banks and test items available for the basic skills of reading, mathematics and language arts. There is a shortage of item banks in more difficult-to-measure subjects such as social studies, vocational education and literature. There is a surprisingly small number of item banks in science and health. And there are almost no item banks for more subjective areas such as art and music. Yet, even though there are areas where more items are needed, the construction of item banks has not been held back by these shortages.

What may be missing is a science of item banking. Most educational innovations have had the support of vocal, highly visible proponents who have actively led the growth of the innovation. Item banking has not had such proponents. This in itself is not bad; certainly item banking has not been oversold or inappropriately used as sometimes happens with better publicized innovations. But at the same time, this low-key approach has prevented item banking from capturing the attention and devoted effort of large numbers of educational measurement specialists. The kind of discussion which took place about behavioral objectives, criterion referenced testing, and other technical innovations has not taken place with item banking. Thus the "science" of item banking remains in its infancy. There are few articles about item banking in the literature, and most educators would be unable to name the leaders in the field of item banking. There are just not many people devoted to item banking as an area of professional interest.

While the technical state of the art has not progressed greatly, it is possible to make some observations about the current status of item banking. One important observation is that a good test item is a relatively treasured commodity. The kind of exchange of test items among groups that would facilitate item banking is not generally taking place. The largest repositories of good test items--the test publishing companies and the statewide assessment programs--have no mechanism and, usually, no inclination to make their items widely available. This has resulted in a sort of Gresham's Law for test items--bad test items drive out the good ones. Measurement professionals attempt to horde good test items; the items that are easiest to obtain for item banks are generally not the best. This should not be surprising; good test items are usually the result of work by highly competent item writers and rigorous review and pilot testing. Since these activities cost money, most item sources are loathe to give away valuable items. (As an aside, many of the groups which have large collections of sound items have already developed item banks for their internal use.)

The shortage of good items may be, in part, the cause of the growth of the "itemless item bank." A few groups have begun development of generalized systems, often computer-based, which have the framework for classifying and retrieving items, but which leave it to the user to insert test items. In one sense, these systems are valuable since they will allow the creation of new test items by people who purchase the systems. But it is unfortunate that these itemless item banks do nothing to get good test items to educators who need them but cannot develop them themselves.

Another observation about the current state of item banking is that the decline in the amount of discussion going on about criterion referenced and minimum competency testing has led to some reduction in the amount of discussion about item banking. There is little question that item banks "missed their chance," so to speak, when there were inadequate numbers of useful banks available during the growth of the CRT/MCT movement during the 1970s. Many educators in the field have by now developed some way of coping with testing requirements that does not rely on an item bank. Thus, some of the impetus for developing large, high quality item banks is gone.

A final observation is that some of the item bank developers may have overestimated the measurement sophistication of the educators who need to use the banks. This lack of sophistication has manifested itself in two ways. First, some item bank developers thought that the quality of an item was not absolutely critical, since even a less than perfect item would provide an advantageous starting point for test development. It may be, however, that educators in the field lack the training necessary to make the needed improvements. Items that start with problems may end up with exactly the same problems. Second, some developers of item banks had a strong criterion referenced orientation which they tried to maintain by organizing their items with elaborate classifications which allowed users to select items by specific content objectives. The result was that large numbers of items were needed in order to cover all of the objectives that might possibly be contained in the bank. This is not the place to debate the value of elaborate item specifications and classification systems, but it does appear that educators in the field might prefer fewer, more general items that are of uniformly high quality. A request for good fourth grade math items is more common than a request for a "two-digit plus three-digit addition with carrying in the ten's place, horizontal format, multiple choice" item.

Nonetheless, item banking remains a concept which receives wide support among educators. The development of test items from scratch seems like a wasteful practice which should be reduced. However, the concept of item banking suffers from a lack of knowledge about how to do it and a shortage of good, available test items to put into banks. Item bank development for general use has not received the widespread attention from professionals required to make item banks proliferate. The status of item banking today is not greatly different than that of two or three years ago. While it is likely that those people who have access to item banks find them useful, there is not yet a mechanism for their widespread development and distribution. Item banking, in 1981, has some distance to go before becoming an integral part of routine educational test development (if, indeed, it ever will be).

II. DESIGN OF THE STUDY

In this section of the report, the procedures which were planned for the study are described. These procedures include the selection of cost accounting categories, operation of the content panels, test review and piloting plans, and analysis plans.

Selection of Accounting Categories

To be able to make statements about the relative cost of the two different test development methods, accurate cost accounting was necessary. One of the first activities in the study was to determine the categories in which costs would need to be maintained. Three categories were assumed to cover all the costs associated with the project; these were NWREL staff costs, content panel costs and materials development costs. Records of these costs were to be maintained in two ways. First, the NWREL testing specialist was to keep a record of the amount of time every person involved in the study (himself, other NWREL professional and support staff, content panel members) spent working on the test development efforts. Second, the NWREL monthly accounting reports were to be reviewed to corroborate the time records maintained by the testing specialist and also to determine the other costs associated with the project (e.g., duplication and graphic art costs).

Operation of Content Panels and Plans for Test Development

Over the past several years, NWREL has worked with a number of teachers, administrators and curriculum specialists who have served as content panel members in a variety of contracted test development activities. Many procedures have been worked out to make content panel meetings as efficient as possible. Those procedures were to be used as a starting point for this study. Indeed, the content blueprints with which the panels were to begin work were drawn from past work NWREL staff have done for the Alaska Statewide Assessment Tests. Both types of panels (item bank and item specification) were to work from the same content blueprint, but the procedures to be used with each panel were different.

The procedures NWREL had used with other content panels were to be followed in working with the two item specification panels (reading and math). Those procedures are as follows:

1. Distribute content blueprint to the panels. No selection among content included in the blueprint will be necessary. (This is perhaps the most unrealistic aspect of the study in that usually panels must select a subset of content to be tested from among all possible content included in the blueprint. Nevertheless, for the

purpose of this study, it was felt that this was not a critical step which needed to be followed and compared.)

2. NWREL staff develop behaviorally stated objectives matching the blueprint entries. Content panels review objectives, which are then modified as necessary.

3. NWREL staff write item specifications from the statements of objectives and develop a sample item for each item specification. Content panels review item specifications and sample items, which are then modified as necessary.

4. NWREL staff write the appropriate number of items from each item specification to complete the 45-item tests. Content panels review test items, which are then modified as necessary.

5. NWREL staff format items into final tests. Content panels review final tests, which are then modified as necessary.

Whereas already developed procedures were to be used with the item specification panels, new procedures were to be developed for working with the item bank panels. Those panels were to receive the same content blueprints as the respective item specification panels. Rather than go through development of behavioral objectives, item specifications and sample items, though, panelists were to go directly to the appropriate sections of NWREL's item bank. (See Appendix A for a description of that item bank.) Nine items from each relevant item bank section (three times as many as were needed for that portion of the final test) were to be selected by the NWREL testing specialist and presented for the panelists' review. Of the nine items with the appropriate item bank descriptor, three of them were to be the items from the item specification panel's test. The other six items were to be drawn randomly from the bank's entries at the appropriate difficulty level. Panelists were to be presented the nine items and told to select the three which they felt should be on a fourth grade test measuring the skill included on the content blueprint; the panelists were to be responsible for choosing the three items that, in their opinion, provided the most useful information related to the skill. When all 45 items were selected, they were to be formatted into a final test by the NWREL testing specialist, reviewed by the content panel and modified as necessary.

Test Review, Piloting and Analysis Plans

The tests developed through the two different methods were to be studied in two ways. First, they were to be administered to a group of beginning fifth grade students so that their empirical psychometric characteristics could be studied. (It was necessary to use fifth grade students because of the timing of the project; the tests were developed for end of fourth grade students, but the study was to end December 31. That was judged to be too early in the year to administer the tests to fourth graders.) Within one subject area (reading or math), each test was to be given to a random half of the students in enough classrooms so that the total sample size was 75-100 students for each test. Item analyses were to be conducted and statistical tests of the differences between test means, standard deviations and reliabilities were to be computed.

The second approach to determining the quality of the tests was to have them reviewed by elementary school teachers and appropriate content specialists. Two review instruments were to be used. The first would contain all the items (but in random order) from the two tests in the same content area, along with a copy of the content blueprint. Reviewers (approximately 50 for the reading test and 50 for the math test) would be asked to match each test item to an entry in the content blueprint. It was assumed that if a test has content validity, it will be obvious which items measure specific parts of the content blueprint. If one test had more items correctly matched, it could be said that it had more content validity. The proportion of items correctly matched to the blueprint would be determined for each test and then a statistical test for the significance of differences between proportions would be performed to see if one type of test has greater content validity than the other type.

In addition, a second group of reviewers (again approximately 50 in each subject area) were to be sent a copy of the two tests and asked to respond to a brief questionnaire. The questionnaire was to include items related to the degree of coverage of the content blueprint, homogeneity of the items measuring each objective and overall quality of the test. Reviewers were to be asked to respond by saying that Test A was better on the given dimension, Test B was better or there was no difference. Chi square tests were to be performed on the results to see if either test was perceived as significantly better than the other on any of the dimensions.

III. CONDUCT OF THE STUDY

This section of the report describes the actual implementation of the study design. In most cases, the study operated as planned. Where discrepancies between the design and its operationalization occurred, they are noted. The sections below describe selection of the content panels and test content, operation of the four content panels, and content review and pilot testing of the four tests.

Selection of Content Panels

In late 1979, a classified advertisement was placed in Portland newspapers soliciting teachers and curriculum specialists to participate in an upcoming item bank project being conducted by NWREL. Over 200 persons responded to the ad and completed a background information sheet regarding their area and grade level of expertise, prior item writing experiences, etc. On the basis of this information, a subgroup of approximately 50 persons was selected for interviews.

These 50 information sheets were retained in NWREL files and used again in the current study. The sheets were reviewed to identify persons whose area of expertise was elementary school reading or math. A total of 28 persons met that criterion. A letter was written to each person describing the current study and asking them to call the principal investigator if they were interested in participating. Nineteen persons called to say they would like to participate. One of the nineteen advised NWREL staff of a friend with the proper qualifications who wanted to volunteer if additional persons were needed. Since the design called for twenty content panel members (four panels with five persons on each), the friend was contacted and hired as well.

Thus the design was fulfilled in that 20 persons were selected to serve as content panel members. When attempts were made to set up initial content panel meetings, however, two individuals could not be reached. The result was that the project proceeded with 18 content panel members rather than 20. The two item specification panels did have five members each, but the two item bank panels had four members each. The design is shown below.

Design for Study of
Two Test Development Strategies

	Subject Area	
	Reading Panels	Math Panels
Item Specifications	Panel A1 (n=5)	Panel A2 (n=5)
Item Bank	Panel B1 (n=4)	Panel B2 (n=4)

Copies of the letters written to potential panelists and those selected to participate are included in Appendix B and C, respectively.

Selection of Test Content

The study directors were able to take advantage of previous work conducted by NWREL's Assessment and Measurement Program by selecting as the tests' content the same content that had been covered in the fourth grade Alaska Statewide Assessment Test. The advantage to this was that (1) "real" content panels had decided that the content was appropriate for fourth grade students and (2) that assured there would be items to match the content in the item bank since the actual assessment test items were in the NWREL bank. Reading comprehension and mathematics computation were selected as the specific content areas because it was estimated that those areas represented the two ends of the continuum in terms of difficulty in developing basic skills tests--reading comprehension being very difficult, math computation being very easy.

While the study design had specified that fifteen topic areas would be selected for each of the tests (reading and math), only nine math areas were included in the Alaska Assessment. Because time and cost comparisons would be made for the reading and math tests, it was decided to limit the reading test to nine areas as well. The original plan of having each test consist of 45 items was maintained, but those items were spread over nine content topics rather than fifteen.

The item specification panels decided how the 45 items would be divided among the nine topics during early meetings. So that comparisons between the item specification and item bank tests would not be confounded by different content emphases, the item bank panels were given the number of items that they had to have for each topic area. Thus, the final tests had the same number of, for example, division problems. The assumption was made that it would have taken the item bank panels the same amount of time as the item specification panels to decide on the content emphases, and so that amount of time was added to the item bank panels' total time spent in test development.

Operation of Item Specification Panels

The reading and math item specification panels met in the afternoon of June 23. Since the orientation to the purpose of the project and description of the tasks were similar for both the math and reading panels, the meeting began with both panels together. The panels were first introduced to the project's objective--the comparison of two procedures for constructing an achievement test.

Both sets of panels were given a list of the nine content areas that were to be covered in each of the tests. The content blueprints given to the panels were:

Reading Comprehension

1. Use Stated Facts to Answer Question
2. Determine Sequence of Events in a Passage
3. Identify Cause and Effect Relationships
4. Determine Object of a Referent
5. Identify Main Ideas
6. Identify Conclusions and Supporting Facts
7. Distinguish Fact, Opinion, Fiction, Non-Fiction
8. Infer Formulations (Conclusions, Generalizations, Unstated Assumptions)
9. Generate Solutions, Predict Outcomes, Apply Ideas to New situations

Mathematics Computation

1. Add Whole Numbers Without Regrouping
2. Add Whole Numbers With Regrouping
3. Add Decimals
4. Subtract Whole Numbers Without Regrouping
5. Subtract Whole Numbers With Regrouping
6. Subtract Decimals
7. Multiply Whole Numbers Without Carrying
8. Multiply Whole Numbers With Carrying
9. Divide Whole Numbers Without Remainder

The panels then discussed the general test description information that applied to all the tests in the study. The panels were told that the tests that they were designing were to have 45 items and they could distribute the items among the content areas as they chose. The only limitation was that there needed to be at least three items per content area; that was determined to be the minimum number of items related to a topic one could have on a test and still allow a score to be reported for that topic. The panels were also told that the content areas were broad and they had the option to concentrate their measurement on a limited area within the content area. In other words, if the math panel chose to use only three-digit numbers in the "Add Whole Numbers With Regrouping" section, they should make that limitation clear to the NWREL testing specialist who would write the items and also to the people who would be scoring the test and using the limitation. The panels also discussed the range of difficulty that the items should cover. This involved establishing that the test was to be viewed as criterion-reference and should be based on skills that students should have at the end of the fourth grade. This then eliminated the need for very easy items and very difficult items for students at the extremes of the ability distribution within the grade.

The panels were then instructed that the remainder of the meeting should be spent reading over the content blueprint and defining these areas so that item specifications could be written for them. The panels were given a brief introduction and the description of an item specification as presented in the Introduction section of this report.

After these introductory comments, the two panels separated and spent the rest of the meeting working as individual panels.

Reading Item Specification Panel Activities. In retrospect, the reading item specification panel had the most complicated task of the four panels involved in this study. This is reflected in the amount of time that it took to construct their test. It was difficult to separate the writing of item specifications from what the final test was going to look like. If one considers the combination of passage length, passage topics and item types, it can be seen that there are a large number of item specifications that could have been written for the item specification panel.

The rest of the first introductory meeting was spent reading over the content areas and providing more information so that item specifications could be written. The panel also took this time to think about some other test specification information such as passage length, paragraph topics, and other limitations. The notes from that first meeting were typed and used to write the item specifications.

Based on the notes from the first meeting, item specifications were written in the Popham format. The item specification format calls for a sample item to be included. In this study, no original sample items were written. Based on the notes from the first meeting, it appeared that 19 separate item specifications needed to be written. Writing these passages would be very time consuming especially when the main purpose of the initial draft of an item specification was to see whether the item format was close to what the panel wanted. Therefore, rather than write the 19 passages needed for the item specifications, it was decided to go through available standardized tests and choose passages that would serve as sample items in the item specifications. Since this study was to look at the procedures as they would probably occur in the field, this strategy seemed appropriate.

At the second meeting, the panel was given the item specifications and asked to determine whether items written to these specifications would be satisfactory for the test. The panel was also asked to begin thinking about the number of items that should be written to each specification keeping in mind the limitations presented at the first meeting (at least three items per content area, a total of 45 items, and items appropriate for end of fourth grade students.

The use of sample items from standardized tests did save time in the writing and also brought up other points that the panel decided should be limitations or criteria for the items. For example, after they saw the sample items, they decided that "None of the above" and "All of the above" should be avoided. Also, they decided that the questions should all be written in the affirmative; negative statements should be avoided because they often confuse students. It was also decided that the student would be asked to choose the "best" answer in the general instructions to the test. Thus, a response could be close to correct and better than the alternatives, but not the completely correct answer. Also, topics and styles that should be omitted or avoided if possible were identified.

This second meeting was quite long (6 hours). Panel members critiqued the item specifications, allocated the number of items per content area, the number of items per item specification and decided the number and kinds of passages needed for the test. The numbers of items per content area that the panel chose were:

Content Area		# of Item Specs	Number of Items	
1	(Stated Facts)	4	4	
2	(Sequence)	4	6	
3	(Cause and Effect)	2	6	
4	(Object of Referent)	1	3	
5	(Main Idea)	3	6	
6	(Conclusion)	2	6	
7	(Fact, Opinion, Fiction, Non-Fiction)	2	6	(4 for Fact or Opinion; 2 for Fiction and Non-Fiction)
8	(Formulation)	3	3	(1 each for Conclusions, Generalizations, Unstated Assumptions)
9	(Solution, Outcome, Application)	3	5	(2 for Generate Solutions; 2 for Predict Outcomes; 1 for Apply Ideas to New Situations)

The panel also decided on the total number of passages to include in the test and the number of items to follow each passage. They asked that the test contain nine passages:

3 Shorter Passages — 5 to 6 short, simple sentences or 3 complex sentences

4 Medium Passages -- 10 to 12 sentences

2 Longer Passages — up to 4 paragraphs

They also decided on the number of items to follow each passage:

Shorter Passages — 3 to 4 items
Medium Passage -- 4 to 6 items
Longer Passages — 5 to 8 items

Because of some problems with project staff schedules, and attempting to keep to timelines, it was necessary for another NWREL staff member to become involved in the study at this time. A NWREL professional staff member with item writing experience wrote the first draft of the items that appear on the test. She used a matrix of the number of items per content area plus the notes from the second meeting of the panel to write the items.

At the third meeting of the reading item specification panel, panelists reacted to the items that had been written. They had the option of rewriting any of the passages or questions or asking that they be rewritten for their approval. Both of these options were taken. The evaluation specialist had written 55 items to 9 passages. The panel:

1. Rewrote 5 of the 9 passages
2. Approved 35 items
3. Eliminated 11 items
4. Rewrote 9 items
5. Asked that 2 additional items be written

A fourth meeting of the reading item specification panel was held to review the items after they had been rewritten from the third meeting. The panel agreed that they could have probably met even a fifth or sixth time before they would have been perfectly satisfied with the test. This panel seemed to have more personal identification with the test passages than did the later item bank panel. This could be due to the limitations imposed on the item bank panel (to be discussed later in the report). The item specification panel was very conscientious in their endeavor to put together a good test but it seems that one of the disadvantages to the item specification technique used here was that a deadline date was not chosing something that should be a part of future item specification panel test development activities.

After the final meeting of the item specification panel the test was formatted and reproduced. Appendix D contains the reading item specifications and Appendix E contains the test developed by this panel.

Math Item Specification Panel Activities. After the initial combined meeting of the item specification panels on June 23, the math item specification panel met separately from the reading panel to work on providing the necessary information (including deciding how many items to include for each section of the test content) so that the item specificatons for the math test could be written.

From this information, 248 item specifications were produced. The reason that so many item specifications were produced is that there are a variety of ways that the content area "Add Whole Numbers Without Regrouping," for example, can be exemplified. There are numerous variations of one-, two- and three-digit problems each demanding its own item specification. The actual production of these item specifications did not take a long time, however, because a word processor was used. Also, math problems that needed to be written for the sample item in the item specification do not take a long time to construct.

The 248 final item specifications were then given to the math panel at a second meeting for them to go through and decide which should have items written to them to appear on the test. The panel chose 45 item specifications to be written for the final test. The number of item specifications and items per content area were:

Content Area	# of Item Specs	# of Items
1. Add Whole Numbers Without Regrouping	3	3
2. Add Whole Numbers With Regrouping	4	4
3. Decimals	5	5
4. Subtract Whole Numbers Without Regrouping	3	3
5. Subtract Whole Numbers With Regrouping	6	6
6. Decimals	6	6
7. Multiplying Whole Numbers Without Carrying	7	7
8. Multiplying Whole Numbers With Carrying	6	6
9. Divide Whole Numbers Without Remainders	5	5

The next step was to write items to these item specifications. At the suggestion of the content panel, the sample items which had accompanied the item specifications were used for the test. It was also decided at this meeting that all problems involving more than one decimal place value, that is two decimal places, should be written in a dollars and cents format. The panel thought that this was more familiar and one of the textbooks that they were using for a reference always used this format.

From the item specifications, 45 items were written. This happened after the meeting, but this panel also had to decide on some of the parameters of the test format and this caused some discussion. First, there had to be agreement on whether the decimal points should be lined up in the distractors. There was not much problem in coming to an agreement that they should.

The discussion that took the longest was whether there should be commas in the distractors. There was no disagreement on the use of commas where there were more than four digits in the number, but the problem arose where there were only four digits—should it be "2,119" or "2119"? The problem occurred in items where there were both four and five digit distractors. Even when it was finally decided that the commas should appear in the four-digit numbers, there was some concern because to remain consistent the panel decided that four-digit numbers in the problems themselves should probably have commas. This caused some concern because the panel thought that most students didn't use commas in the problems when they did them. It was finally decided that the items should remain consistent, and so commas were inserted in all numbers of four digits or more.

The panel did not meet for a third time. It was decided at the second meeting that the items were satisfactory and that a prototype of the test would be put together and sent to the panelists by mail. The panel was then told to call in their comments once they saw the items in the final form.

One panel member called with two comments. First, there was some concern on the amount of space between the "$" and the answers in some of the problems. This was caused by lining up the dollar signs. It was decided in these cases to move the dollar sign closer to the numbers in the distractors. The other concern was that perhaps the distractors should be arranged in numerical order from smallest to largest. The fear was that the math test may also be testing for visual discrimination rather than merely mathematical computation skills. This concern seemed warranted but when the items were arranged in this manner, response B was correct a disproportionate amount of time. Since each distractor had been developed to match a particular error students might make, they could not be easily changed to allow for a more balanced distribution of correct responses. The concern then was that the test would have to be written over again as well as the item specifications to avoid this problem. It was decided to leave the items formatted in the original manner (that is, with options arranged more or less randomly, but with each position--A, B, C, D--being correct about the same number of times).

The math item specifications and the final test developed by this method appear in Appendix F and G, respectively.

Operation of Item Bank Panels

Before the item bank panels met, it was necessary to select items from the NWREL item bank for them to look over at their first meeting. One of the first decisions that had to be made that was not part of the original study design was that the number of items for each topic area had to be the same for the item bank test as for the item specification test. The item specificaton panels had been allowed to choose the number of items to cover each topic area. If the item bank panels had been given the same option, there was a concern that the relative content emphases of the two tests might differ enough to make a comparison of the two tests very difficult. Therefore, it was decided to present the two item bank panels with the number of items for each topic area that they had to cover.

The original design of the study required that the item bank panels be given three times the number of items they needed for each of the topic areas and that one-third of these items be the items that the item specification panels had originated. That requirement was followed, as explained below.

First to be selected from the NWREL item bank were the math items. Table 1 shows the number of items that were needed for each content area, the number that came from the item specification panels, the number needed to be chosen from the item bank and the number of items that were available to be chosen from the bank.

Table 1
Number of Items Selected for Review
from Math Item Bank

Content Area	Total Number of Items Needed	Number of Items from IS Panel	Number of Items from Item Bank	Number Available from Item Bank
1	9	3	6	80
2	12	4	8	85
3	14	5	10	33
4	9	3	6	102
5	18	6	12	56
6	18	6	12	43
7	21	7	14	78
8	18	6	12	69
9	15	5	10	86

The process used to select items from the bank to present to the panel was as follows:

1. Using the item assignment cards, all items in a topic area were numbered in sequence.

2. A table of random numbers was used to randomly select the numbers of the items to be chosen from the bank.

3. The items were selected from the bank.

4. An item was identified as "Not Usable" if:

   a. it was missing from the item bank[4];
   b. it was a story problem rather than a straightforward computation item;
   c. it was not multiple choice.

5. The steps were repeated until the number of items needed from the bank was reached.

The process took three iterations before all the categories had the correct number of items.

The process for drawing items from the reading bank was similar to that used for the math panel but had a few more complications. First, the following information that was similar to that of the math panel was needed.

Table 2
Number of Items Selected for Review
from Reading Item Bank

Content Area	Total Number of Items Needed	Number of Items from IS Panel	Number of Items from Item Bank	Number Available from Item Bank
1	12	4	8	292
2	18	6	12	34
3	18	6	12	23
4	9	3	6	55
5	18	6	12	222
6	18	6	12	73
7	18	6	12	467
8	9	3	6	226
9	15	5	10	26

---

[4] There was a problem in that some of the math item sheets are missing from the NWREL item bank. So it happened, occasionally, that there was an item assignment card selected but its matching item could not be found.

The first problem here was not knowing from the item bank index the level at which the item was written or the difficulty level of the item. The only way to find this information was to refer to the actual item. Some of the items in the bank have a suggested grade level, but most do not so that decision was left to the discretion of the NWREL testing specialist. Items were eliminated only if they were obviously too easy or difficult for fourth grade students.

Following the same procedure as used for the math items did not seem to be a good plan. It was obviously going to take many iterations of the process. It was decided that in the long run it might be easier and less frustrating if the items from the bank were reviewed quickly and those items that were obviously not usable eliminated. This took a good deal of time and the following list shows the number of items in each category that were finally considered available for the test.

Table 3
Final Number of Reading Items
Finally Available for Review

Content Area	Number of Items Needed from Bank	Number of Items Indexed in the Bank	Number of Items Considered Usable
1	8	292	81
2	12	34	16
3	12	23	8
4	6	55	15
5	12	222	36
6	12	73	23
7	12	467	262
8	6	226	39
9	10	26	9

It is obvious that this editing procedure greatly reduced the number of available items. The major reason for eliminating items was that they were not at an appropriate difficulty level for the end-of-four-grade tests being developed.

After the item bank had been reduced, the same procedures that had been used for the math items were used to select reading items. Although it appears initially that the reading item bank panel had the same number of items to review as did the math item bank panel, it should be remembered that when a reading item is referenced in the item bank, a question is selected and then a reading passage accompanies the item. Some of these passages are quite long.

Therefore, the packet of materials that was developed for the reading item bank panel was very thick and contained a large amount of reading.

Introductory Meeting. Both the reading and math item bank panels met together for the first time on the evening of August 18. They were given an introduction to the project similar to that given to the item specification panels in June. One difference was that the item bank panels were not given the freedom to choose the number of items per content area. The rationale for this limitation was explained to the panels.

After the initial introduction to the project, the two panels were then asked to work separately and to start their selection of the items that were to appear on their tests.

Reading Item Bank Panel Activities. The item bank panel for the reading comprehension test met the remainder of the evening and looked through the items from the item bank. They also met a second evening to finish the task.

The product of these two meetings was 45 items. However, the selection of 26 passages was necessary in order to get these 45 items. There had been no time limit set for the test but there was a consensus among panels that the test should not take longer than 45 minutes. For the students to read through 26 passages and to answer 45 items would certainly have taken much longer. Therefore, the item bank panel was asked to meet again to see if they could reach agreement on a remedy for the problem.

After the item bank panel met for the third time, they had reduced the number of passages considerably. The panel still had a problem, however. They had reduced the number of passages but in doing so they were short on some of the content areas needed to be covered by the test. The only solution seemed to be for new items to be written to the passages that were chosen for final inclusion in the test. This was not in the initial plans of the study, but was an unforeseen necessity. It also provides, though, some useful information about the problems that can be avoided in future applications in the use of an item bank.

One of the panel members volunteered to write the remaining 16 items. It was decided that this should be the procedure since the panel member knew what the panel was looking for in the needed items. It should be noted that this perhaps violates what could be termed a "pure" item bank procedure by employing an item specification method. There was no alternative that was feasible to remain within timelines. It does also, however, highlight a question for future test developers who use item banks--what does one do when the available items are not acceptable?

After the items were written and the test had been drafted, this draft was sent to the panel members for their final approval. Each of the panel members was contacted to see if they had any comments about the quality of the items and the format of the test. The final reading comprehension test constructed by this panel appears in Appendix H.

Math Item Bank Activities. The math item bank panel met the remainder of the evening of the introductory meeting and then met again two nights later to complete their task of choosing items from the item bank for the test. Their job went very smoothly and the only real questions seemed to be whether problems should be in a vertical or horizontal format and the ordering of the test items. The item bank panel put all the decimal problems together in the test which is different from the way they appeared on the content blueprint. This is also different from the ordering that the item specification panel chose.

As with the reading item bank panel, the final math items were put together on the test and sent to the panel members for their approval. This final test appears in Appendix I.

An observation that was made during the construction of the item bank math test by the item bank panel should be noted here. The item specification panel was forced to formalize their choices of the types and content of items into item specifications. The lack of the formal item specification does not mean that the item bank panel did not approach the construction of the test in an orderly fashion. In fact, when observing the panel at work, it was obvious that approximately the same procedures were used by both types of panels. For example, when the math item bank panel looked at a particular content area on the blueprint, it was common for them to vocally identify what kind of items they thought were needed and then they would look through the list from the item bank and choose the item that most closely matched this need. The item specification panel needed another step for the item writer to write the formal item specification. The problem with the item bank method was that the panel could not always find exactly the kind of item they were searching for and most often had to settle for an item that was close. This could be a result of the item bank procedure but could also be a result of the plan of this study where the item bank panel was limited in the number of items it was given to select from.

Content Review of Tests

Although the major test of the two test development strategies was an empirical one (determining their psychometric characteristics after a pilot administration), it was felt important to learn teachers' and curriculum specialists' perceptions of the tests to see if there were differences in how the tests were viewed. Two different questionnaires were developed to gather information about the tests. One questionnaire attached two intact tests (either the two math tests or the two reading tests) and asked respondents whether they thought one test was better than the other on several dimensions. The second questionnaire contained the items from the two math tests or two reading tests intermingled so that it was impossible to tell from which test the items had come. Respondents were asked to match each item with one of the nine topic areas which formed the framework for the test.

Working through the evaluation offices of four school districts (Beaverton, OR; Boulder Valley, CO; Grand Rapids, MI; Haslett, MI), 140 volunteer respondents were identified. Questionnaires were mailed to the district evaluation offices during the week of October 26, 1981. Each evaluation director established his or her own procedures for distributing the questionnaires. Respondents mailed completed questionnaires directly to NWREL in postage paid envelopes, which were provided with the questionnaires.

Copies of the correspondence to the district evaluation offices and the questionnaires used for the content review of the tests are found in Appendix J.

Pilot Testing

A very important determinant of the quality of a test--and, in this case, of test development strategies--is how the test performs when it is administered to a group of students. It was felt that the tests developed as part of this study must be judged empirically by looking at their psychometric quality. To do this, the tests were administered to fifth grade students in the classrooms of eight volunteer teachers from the Central Valley (WA) School District. Testing took place in the first half of November 1981. Fifth graders were selected for testing rather than fourth graders because the tests were designed as end of fourth grade tests.

Within each classroom, approximately half the students answered one test (e.g., the item bank math test) and half the students answered the other test (the item specification math test). The two tests in one subject area were alternated in the pile of tests given to the teacher so that a random half of each class took each test. The tests were identified on their front sheet as being either Form IB or IS. A separate answer sheet was used, and students wrote either IS or IB on their answer sheet to show which test they were answering.

Copies of the answer sheets and teacher directions are provided in Appendix K. The tests have already been presented in earlier appendices.

IV. RESULTS OF THE STUDY

Results from the four different reviews of the tests are included in this section of the report. First, results from the pilot test administration are described. Then, questionnaire results from the teachers and content specialists are provided. Next, an analysis of the time required to develop each test is given. Finally, a similar analysis of the cost of test development for each test is provided. Implications of these results are provided in the final section of this report.

Results from the Pilot Test: Assessment of Psychometric Properties

In mid-November 1981, 16 fifth grade classrooms in the Central Valley (WA) School District participated in the pilot tests. Each math test was taken by 72 students. The item bank reading test was taken by 95 students; the item specification reading test was taken by 96 students. Answer sheets from the students were keypunched, coded as to the test they took, and the card deck submitted to the LERTAP (Laboratory of Educational Research Test Analysis Package) computer program at Bonneville Power Administration's CDC 6500 computer. Mean scores, standard deviations and Hoyt estimates of reliability[5] for the four tests are shown in Table 4 below:

Table 4
Means, Standard Errors and Reliabilities

Type of Test	N	Mean Score	Standard Deviations	Reliability
Item Bank Math	72	32.8	7.92	.91
Item Specification Math	72	36.2	6.56	.87
Item Bank Reading	95	33.0	6.15	.82
Item Specification Reading	96	32.4	7.76	.87

---

[5] Hoyt's ANOVA approach to internal consistency reliability estimates is algebraically equivalent to KR-20, but allows for the differential weighting of test items. Such differential weighting was not used in the tests under investigation here.

Statistical tests--t, F and z tests--conducted to determine the significance of differences means, standard deviations and reliability estimates, respectively, showed the following results:

Table 5
Comparison of Means, Standard Deviations and Reliabilities

Comparison	Means	Standard Deviations	Reliabilities
Math Tests	$p < .01$ (item bank test easier)	Not significant	Not significant
Reading Tests	Not significant	$p < .02$ (item bank test less variable)	Not significant

The comparison among means is not really an indicator of differing quality based on the test development method, but it was performed just to make sure that all of the tests were of an appropriate difficulty for the students tested. The fact that four of the six comparisons were not statistically significant is ample evidence that neither test development strategy produced definitely superior tests, at least from a psychometric standpoint. Even though the reliabilities were not significantly different, it is interesting to note that the math item bank test--the most straightforward of the four tests to develop--had the highest reliability; whereas the reading item bank test--the test that had the least structure in its development (in that the panel first selected too many passages, had to delete some and in so doing made it necessary to write new items, but without benefit of item specifications), had the lowest reliability. In fact, those two reliabilities (.91 and .82) are significantly different from each other at the .05 level.

Related to this difference is the difference between the math and reading item bank tests' standard deviations. The math item bank test has significantly greater variability than does the reading item bank test. Given that result and the fact that high test reliability is dependent, in part, on high variability among test scores, it is not surprising that the item bank test showed a higher reliability than the other test.

A final difference between the tests as a whole that should be noted is that the math test developed using the item bank procedure had a lower mean score than that developed using item specifications. That is, the item bank math test appears to be more difficult. Although the content blueprints for the two procedures were the same, the item bank panel did include decimal problems containing more than two decimal places, the limit set by the item specification panel. No comparative difficulty of subsets of items was done to study if the difficulty could be attributed to a content area, however.

Thus, in summary, it appears that differences in the tests developed through this study are not due so much to the test development method used, but rather to the content tested--at least as far as test reliability is concerned, and at least for the item bank tests. Somewhat surprisingly, the two item specification tests had exactly the same reliability (.87), even though the standard deviations were quite different (but not statistically different).

In addition to looking at overall test statistics, each item in the four tests was also reviewed. The LERTAP results showed how each item on the tests performed by providing p values and point-biserial correlations with total test score for each option on each item. Four criteria used to identify possibly poor test items were selected for special attention, as follows:

1. A correct option with a p value less than .50--Since the tests were supposed to be for end of fourth grade students, it was felt that any item that was answered correctly, by fewer than half the third month fifth graders was suspect.

2. A correct option with a p value less than any distractor's p value--Anytime a test developer sees a correct option with a p value less than a distractor's p value, he or she worries about a mistake in the key. If the item was, in fact, keyed correctly then one has reason to suspect tricky wording in the item that makes a distractor seem like the correct answer.

3. A correct option point-biserial correlation less than .20--There are no established lower limits as to an acceptable point biserial, particularly for criterion referenced items. Several authors, however, suggest that any item with a point-biserial less than .20 deserves scrutiny. Needless to say, the higher the point-biserial, the better the item discrimination.

4. A distractor with a point-biserial greater than zero--This indicates a definite problem since it suggests that students who otherwise performed well on the test tended to select an incorrect option. Distractors such as these usually need to be changed to make them less attractive to high performing students.

Looking at the four tests developed in this study results in the following number of items marked for special review because they met one or more of the above criteria.

Table 6
Number of Items With Poor Item Characteristics

Type of Test	correct p value $< .50$	correct p value $<$ incorrect p value	correct point biserial $< .20$	incorrect point biserial greater than 0.00
Item Bank Math	7	3	3	12
Item Spec. Math	1	0	9	3
Item Bank Reading	4	2	9	10
Item Spec. Reading	4	4	0	3

The only consistent result evident from the table is in the last column. Both item bank tests had substantially more items than the item specification tests in which incorrect distractors had point biserial values greater than 0.00. One explanation for this is that the item specifications gave specific details for what each item's distractors should be; each distractor represented a likely error students would make. Thus, students who knew the skill being tested in an item would not likely select an option where a specific error was in evidence. The item bank distractors, on the other hand, seemed to be more random. In fact, several times the item bank panel (especially the math item bank panel) questioned where a particular distractor had come from, since they could see no logical (but incorrect) explanation for coming up with the answer. Of course, items in item banks should be reviewed (and modified if necessary) before being used on a test so that only items with appropriate distractors are included. The review and modification process should be extended to all items so that only items with good measurement characteristics are included in the pool. The high number of items with poor discrimination in this study is more an indication of a problem with specific items in the NWREL item bank than with the concept of item banking. As noted earlier in the section on the current status of item banks, some of the items offered for inclusion in item banks have not received the technical scrutiny required of, for example, items on standardized tests. It is, therefore, important to separate the current reality from the potential--and results from this study support the belief that time spent to review and modify items placed into item banks is well worth the effort.

The remaining three columns in Table 6 either show no difference between the two test development methods or show differences but in an inconsistent way (e.g., the item specification math test has more items than the item bank math test whose correct option showed a point biserial less than .20, but the item specification reading test has fewer such items than the item bank reading test). These results, then, are basically uninterpretable in terms of determining whether one test development strategy is better than another.

Results from the Content Review: Assessment of Content Validity

It was noted earlier that 140 volunteer respondents in four school districts were identified to participate in this study. Of the 140 questionnaires distributed, 70 (50 percent) were returned to NWREL. The number of each type of questionnaire that were returned is shown in Table 7.

Table 7
Questionnaire Respondents

	"Is one test better than the other?"	"Match the items to the content listing"
Reading	14	12
Math	27	17

The small number of responses limited the researcher's ability to apply statistical tests, but the questionnaire responses were tallied and are presented below. In all comparisons, the null hypothesis was that one test would not be consistently preferred over the other test.

Table 8 shows the responses to the first questionnaire for the reading tests. The column on the left repeats the questions as they appeared on the questionnaire.

Table 8
Questionnaire 1 Responses for Reading Tests
(N = 14)

Questionnaire Item	Item Bank Test Was Preferred	Item Spec Test Was Preferred	No Difference
1. Look at the nine topic areas which were to be covered on the tests. Which test do you think does a more complete job of covering this content?	3	5	6
2. Do you think one test's format (the way the questions are arranged and the order in which they appear) is better than the other?	4	5	5
3. These tests are both designed for end-of-fourth-grade pupils. Do you think one test is more appropriate for these students than the other?	4	6	4
4. Do you think there is a difference in terms of the quality of the reading passages included in the tests?	6	6	2
5. As an overall judgment, do you think one test is better than the other?	6	4	3

Table 9 shows the same results for the math tests.

Table 9
Questionnaire 1 Responses for Math Tests
(N = 27)

Questionnaire Item	Item Bank Test Was Preferred	Item Spec Test Was Preferred	No Difference
1. Look at the nine topic areas which were to be covered on the tests. Which test do you think does a more complete job of covering this content?	7	4	16
2. Do you think one test's format (the way the questions are arranged and the order in which they appear) is better than the other?	9	12	6
3. These tests are both designed for end-of-fourth-grade pupils. Do you think one test is more appropriate for these students than the other?	6	8	13
4. As an overall judgment, do you think one test is better than the other?	7	6	13

Chi square tests were conducted to determine whether those respondents who expressed a preference consistently chose one test over the other. The small number of responses in each cell and the fact that they were so evenly divided between the two tests contributed to extremely small chi square values. Intuitively, one can look at the tables and see that there was not a consistent preference for the item bank test or the item specification test in either reading or math. The fact that fewer than the desired number of questionnaires were distributed and returned is a flaw, however, and causes less reliance to be placed on these results than might have been hoped.

Table 10 shows the results from Questionnaire 2. The table entries indicate the percent of items on each test correctly matched to their appropriate topic on the topic listing.

Table 10
Questionnaire 2 Results:
Percent of Items Incorrectly Matched to Content Listing

	Item Bank	Item Specification
Reading Test (n=12)	32.41%	33.33%
Math Test (n=17)	5.75%	6.54%

These results are interesting in showing the relative content validity of the reading and math tests. The low percentage of math items incorrectly matched to the content listing allow the inference to be made that both math tests had high content validity. The fact that about a third of the items on each reading test were incorrectly matched is cause for concern, however. If the reading tests were ever going to be administered for a real testing purpose, they would need to be refined further so that one could be more assured of their content validity.

Again, the small number of responses limit the validity of statistical tests; but t tests performed to determine the significance of differences between proportions showed that there was no significant difference between the two tests in one subject area. Thus, responses from both questionnaires, while small in number, corroborate results from pilot tests, which showed no significant differences between the item bank tests and the item specification tests.

Results from the Time and Cost Analyses

The NWREL testing specialist maintained a log of all time spent on the test development activities by NWREL professional and support staff and the content panel members. The number of hours devoted to developing each of the four tests is shown in Table 11.

Table 11
Hours Spent Developing Tests

Test	Professional Staff	Support Staff	Content Panelists	Total
Item Bank Math	12	15	30*	57
Item Spec. Math	27	33	59	119
Item Bank Reading	18	25	56*	99
Item Spec. Reading	42	15	76	133

*These totals were adjusted to add the time of a fifth panel member.

As is apparent from the table, the item specification tests took substantially longer to develop than did the item bank tests. Further, the reading tests took longer to develop than did the math tests; but even the item bank reading test took less time to develop than the item specification math test. It took about twice as long to develop the item specification math test as it did to develop the item bank math test.

That ratio did not hold up for the reading tests, but it should be remembered that 16 new items were written for the item bank reading test. Another point to remember is that although the total amount of time for the item specification test was not twice that required for the item bank reading test, that was the difference in the amount of NWREL professional time spent on the two tests. Since, in this particular test development activity, the NWREL professional staff time was the most expensive commodity, it made the cost of the item bank reading test just about half that of the item specification reading test.

It turned out that the difference in costs of the tests could be totally attributed to the amount of staff and panelists' time spent developing them. Originally, it was expected that reproduction and graphic art charges would be greater for the item specification tests, but that turned out not to be the case. Those costs were just about equal for the two types of tests (item bank and item specifications).

Thus, in summary, it can be said that each item specification test cost just about twice as much to develop as its respective item bank test. Since the test statistics and questionnaire responses showed little or no difference between the two types of tests, it can be said that the use of item banks does seem beneficial and should be pursued for the efficient development of high quality tests. But this conclusion is tempered by several things that were learned about working with item banks which are discussed in the final section of this report.

V. IMPLICATIONS FOR FUTURE TEST DEVELOPMENT EFFORTS

This study was initiated as a comparison of two test construction techniques--the item specification and the item bank procedures. The initial plans of the study often make the two techniques sound completely different. In reality, they cannot be dichotomized quite as easily. One outcome of the study is an analysis of actually using an item bank. There is very little written about the actual steps needed to use an item bank. Therefore, we were exploring somewhat new areas in this study.

Another outcome of this study is the development of ideas regarding construction of an item bank so that it is useful. This study used the NWREL item bank. It should be viewed then as a comparison of that particular item bank's use and not of item banks in general. Many of the steps and procedures would not apply to other item banks. However, the information gathered on using an item bank will apply to other item banks as will the information on how to best set up access to the bank so that it is most useful.

Three problems were encountered in using this item bank and anyone considering constructing, adopting, or adapting an item bank should consider ensuring that these issues are addressed satisfactorily in their own work.

1. There Should be Some Reading Level Indicator for Reading Passages

From the information cards used to access items in the NWREL item bank, there was not enough information to determine the reading level of the passages. Items had to be selected and actually read to get an idea of the reading level. Items could be for a passage consisting of three simple sentences or a quote from Chaucer. The bank was originally designed this way because it was believed that particular types of items would be identified with certain grades by individual users of the bank. Therefore, no problem of identifying the reading level of the passages was anticipated.

Finding a reading level indicator may be a problem in itself, though. There exist a number of algorithms for obtaining rough estimates of reading level. For example, algorithms for three such tests (Bormuth, Fry and Flesch) were applied to the reading passages used in the two reading comprehension tests in the current study. The results of this analysis appear in Table 12.

Table 12
Estimated Reading Level of Passages

Passage	Bormuth Grade Level	Fry Grade Level	Flesch Reading Ease*	Item Spec Test	Item Bank Test
Piggallo	1.9	1	94.8	x	
Jan's Birthday	2.1	2	90.6	x	
Airplanes (1)	5.8	10	56.4	x	
Airplanes (2)	4.2	6	75.3		x
Izzy	3.1	3	87.1	x	
Father's Day	3.0	5	77.2	x	x
Panning Gold	3.2	4	87.9	x	
Chris	2.4	1	93.6	x	
Whales	3.7	4	81.2	x	
John Muir	4.1	5	81.1	x	x
Sugar Beets	3.3	4	85.7		x
North Star King	2.6	3	97		x
Appleseed	7.5	8	68.6		x
Cat Poem	3.9	5	83.8		x
Mohawk River	5.1	7	76.5		x
Archimedes	4.7	6	76.4		x

*The higher the number, the easier the reading passage.

There are definite problems with using any of these indices of reading grade level. The most obvious problem is that there is only minimal agreement on the grade level of the given passage. One explanation in the current application of the formulae is that most of the passages are too short to get a valid indicator of the reading level from the algorithms. Since the algorithms are based on sentence length and number of syllables in the words, short passages can change reading level very easily simply by the addition of a few multisyllable words. A good example is the passage on airplanes, where the item specification panel rewrote the passage slightly. Looking at the reading levels, though, it appears that they changed it radically. The passages do not really differ in difficulty as much as the indicators demonstrate.

Nonetheless, the reading level indicators do at least give a relative scale on which to compare the reading passages. The passages are ranked just about the same by all three methods. It would be suggested, therefore, that for an item bank there be some common algorithm used for all passages so that the user has some information about the relative reading level. Additional information such as teacher ratings of the reading level might also be helpful.

2. Reading Passages Should Offer Many Items to Choose From

One of the problems that the reading item bank panel had was that even though a passage was considered good and the item that was presented as acceptable, there were often only one or two items associated with that passage. This resulted in having many passages to read with only a few items following each reading selection.

It is recommended that as many items and item types be included with a passage in the bank as possible. This will avoid the problem that the item bank panel met.

3. Information Should be Available on the Difficulty of the Items

When an item was selected from the bank, there was not any way of knowing the difficulty of the item just by looking at the descriptive information provided for the item. It could have been a very easy item or a very difficult item for a fourth grader. There was no way to know the difficulty of the item except by actually pulling the item from the bank and studying it.

This is a problem that is encountered by every item bank. One possible solution is the use of item response theory estimates of difficulty. These estimates range from the three parameter models to the simpler Rasch one-parameter model. Several item banks have, in fact, adopted this approach--usually using Rasch calibrations as the way to determine item difficulty. There is study needed in these solutions, however. Even though they claim to be sample free, thereby promising that calibrations obtained through administration of a test to one group will hold up when the items are administered to a different group, there is accumulating evidence that this is not true. If that assumption does not hold, then the use of such calibrations with items in an item bank is not particularly helpful since the entire purpose for an item bank is to be able to use the same items over and over again, with both the same and different students.

Still, this study made it very evident that test development using an item bank would be even more efficient if item difficulties could be associated with individual items in the bank. This is a research area which has received much attention, but where even more is needed.

Summary

The purpose of this study was to discover whether the psychometric quality of objective referenced tests, developed using an item bank, suffered when compared to those developed using the more traditional approach of item specifications and sample items. Both the empirical results from pilot testing and subjective judgments from teacher reviewers indicated there was little or no difference between the tests developed using these two methods. When it is realized that the item specification tests cost about twice as much to develop as the item bank tests, the optimism associated with using item banks seems warranted.

Because this was one of the few studies which documented the use of an item bank to develop tests, it was important to note characteristics which would make item banks even more advantageous. Three recommendations were made which the researchers believe would make item banks easier to use and would make their use even more cost effective.

The use of an item bank to develop a math computation test was quite straightforward and would have resulted in an even better test had more attention been given to eliminating items with poor measurement characteristics (especially those whose distractors do not represent logical errors). Though not as simple, a task such as developing a math computation test, even developing a reading comprehension test, was a less time-consuming task when an item bank was used as a starting point.

The use of item banks in this study did not proceed without complications. The reading panel selected too many passages at first, and then ended up having to write supplementary items. The math panel, intuitively, went through an informal item specification activity. So, in truth, the procedure wasn't "pure." But if the procedure had been pure, there is no indication that the empirical results would have been different and there is every indication that the cost differential would have been even greater. So the conclusion can only be that the use of item banks--even as just a starting point in test development--is a procedure to be recommended.

Appendix A

Northwest
Regional
Educational
Laboratory

PART 4
Item Bank Classification Structures and Procedures

Michael D. Hiscox
Evelyn J. Brzezinski

A Guide to ITEM BANKING in Education

Copyright © Northwest Regional Educational Laboratory, 1980

ITEM BANK CLASSIFICATION STRUCTURES AND PROCEDURES

This document is one of a series of preliminary papers which led to the development of a comprehensive basic skills item bank with NWREL's Assessment and Measurement Program developed for the Shasta County Cooperative Assessment Support System. The system is centered around an extensive collection of test items; these items will allow educators within the County to construct an unlimited number of objective based tests. A method for organizing and accessing the item collection to allow efficient, beneficial use by educational practitioners is presented below. It is not expected that everyone will regard this classification as ideal, and the developers of the collection are interested in receiving comments regarding problems with the system. However, the work of classifying the 20,000 items of the NWREL collection went quite well. The use of the classification scheme in the future will determine the utility of its structure, but to this point, it appears to have some merit.

Item Collection Classification Structures

In the previous section of this paper, a point was made about the multitude of tests being produced every day. It is obvious that this multitude of tests leads to an even greater multitude of test items; an estimate of several million new test items per year would not be overstated.

Clearly, there is no shortage of test items. Why, then, would anyone ever need to write a new test item? There are several reasons:

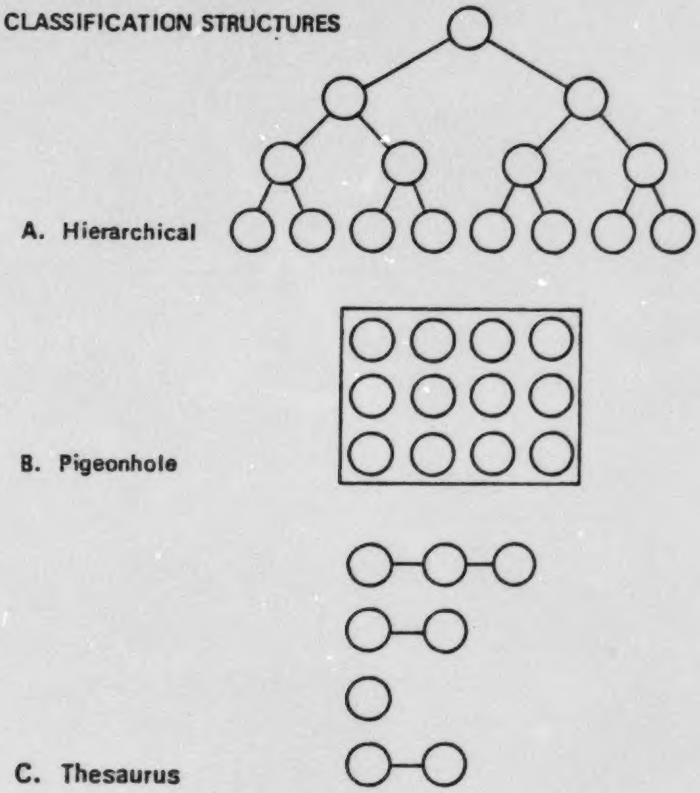
1. The items are not accessible to potential users.
2. The items are of poor technical quality.
3. The items are not retrieved easily from the item collections.
4. The items are not matched to clear indications of what they measure.

While criteria 1 and 2 eliminate a vast number of items from consideration, there are still a very large number of good items available; remember, we started with millions. What prevents the effective use of items written by others is the lack of a method to consistently provide the test constructor with the right item--an item that measures the right skill at the right difficulty level with an appropriate stimulus, response format, and scoring technique. At the present time, there are no examples of "the perfect method." Indeed, it is probable that no classification structure will ever be perfect for all situations. However, the difficulty of using existing classification schemes clearly has been a problem in using item collections effectively.

a. Classification Schemes. Existing item collections seem to fit one of the following categories: they have no sophisticated classification scheme (being grouped only into loosely defined, broad areas); they have a hierarchical classification structure; or they have what might be called, rather inelegantly, a pigeonhole classification structure. Each of these three structures has advantages and disadvantages, which will be described in this section, and each was considered for the NWREL item collection.

This section will also describe an additional structure which has not been commonly used for classifying item collections. This structure will be termed thesaurus classification. While not commonly associated with item collections, the technique is used in various information retrieval systems including the ERIC (Educational Resources Information Center) retrieval system. Individual elements of this structure have also been used in a few other item banks.

The figure below shows a pictorial representation of each of the three classification structures. It may be helpful to recall these simplistic representations as you read about their advantages and disadvantages.


POSSIBLE CLASSIFICATION STRUCTURES

A. Hierarchical

B. Pigeonhole

C. Thesaurus

Some of the structures used with item collections are so loose as to give no structure at all. The items are collected in broad content categories such as reading comprehension or phonetic analysis. But they are not then separated into individual skills. The biggest advantage to this very loose method is the considerable savings of time and effort from not developing a classification scheme and categorizing all of the items. The method, however, is not suitable when it is necessary to retrieve items based on a need to measure attainment of a specific skill.

For example, one would not be able to retrieve an item measuring a student's skill in recognizing consonant blend "CL" except by examining and rejecting virtually the complete set of phonetic analysis items. Since most users of large educational item banks will want to obtain items to measure specific skills, it was considered necessary to have a structured classification system.

b. Hierarchical Classification. Items can also be placed into a hierarchical classification. In this type of system, a content group, like phonetic analysis, is subdivided into individual skills using an outline format. Consider the following abbreviated example:

I. Word Identification Skills
    A. Phonetic Analysis
        1. Consonants
            a. Initial Blends
                1. Initial Blend "CL"
                    a. Oral Word-Word Recognition Item
                        1. Multiple Choice Item

The item:

> "Listen to the word I say...CLOCK. Circle the word in the row that begins with the sound most like the beginning sound of the word CLOCK."
>
> a. class   b. plank   c. corn   d. rock

fits this section of the hierarchy. Hierarchical classifications have some substantial benefits. The classification can define the item completely; that is, one can tell what type of item will be retrieved without seeing the actual item. The classification is expandable by adding more layers to the outline. But the user is also given the chance to use other levels than the final one. Thus, the user could retrieve all initial blend "CL" items regardless of type, or could see all initial blend items, simply by choosing a different level of the hierarchy. Finally, the creation of the hierarchy does not depend on the items collected; the structure comes from the outline description of what topics should be covered. This is useful for identifying topics where the item collection is not comprehensive.

Against these benefits are some significant disadvantages. One is the complexity of the structure. Our example outline was already complex and it only defined one set of items; for example, a full Word Identification Skills structure would be over fifty times as large. The complexity of the structure leads to complicated item coding systems. Our example item might be coded I.A.1.a.1.a.1, an awkward code to handle. Also, the code I.A.1.a.1.a.1 has no intrinsic meaning; it only has meaning when cross-referenced to the hierarchy.

But these minor problems are accompanied by two major problems. First, a hierarchy has no good way to handle items that cover multiple skills; for example, a mathematics problem requiring addition first and then subtraction. Secondly, and more problematic, the hierarchies only work for a few content areas. Over many years, useful hierarchies have been developed for phonetic analysis, math computation, and some other content areas. But some other areas, such as composition or reading comprehension, are not yet, and may never become, hierarchical in nature. Whereas "addition" can be divided into with and without regrouping, horizontal and vertical format, the number of digits involved and so on, "judging the author's purpose" divides less well. This is not to imply that reading comprehension is not as worthy of a hierarchical structure as math computation, only that such hierarchies are not yet established in detail and that they would be time consuming to develop and validate. In short, the hierarchical classification has substantial problems. Better alternatives should be sought for newly developed large item collections.

c. Pigeonhole Classification. Some item collections have been successful in using classification systems where all of the categories are arranged in advance (as with the hierarchical structure) and where (unlike the hierarchical structure) the categories are considered to be at the same level and of the same scope. The structure can best be visualized by picturing one of the old sets of mail sorting pigeonholes. Every pigeonhole is given a name (in this case, a description of content) that may or may not be related to the other pigeonhole names. The names are assigned basically for convenience, rather than for some theoretical reason. In fact, the main criterion may be having one pigeonhole collect as many items as the others.

The classification structure is analogous to those post office pigeonholes. At the post office, sorting names assigned to the pigeonholes might include First Street, Second Street, Third Street, Fourth Street, West Side, Suburban, Other County, In State, National and Europe. There is clearly no parallel between First Street and Europe save that this structure might lead to somewhat the same number of envelopes going into each pigeonhole; the postal worker looking for a single letter would find it about as quickly in one pigeonhole as another. Imagine the situation if the pigeonhole names were changed to Europe, North America, South America, Africa, Asia, Australia and

Antarctica. While the parallel structure is theoretically pleasing, the poor person assigned to North America would have an impossible amount of work while the person assigned Antarctica would rarely need to come in.

Notice also that each of the pigeonholes can have a reasonable name, as opposed to a code like I.A.1.a.1.a.1. The overall classification scheme would read like a list rather than a giant outline, and the assigned names would have some intrinsic meaning.

While pigeonhole classifications cannot identify and code items in minute detail, they are in many ways more practical than hierarchical classifications. And although they are not theoretically as satisfying, pigeonhole classifications are easier to set up and use.

There are, in fact, only three major problems with this type of structure. The first problem is obvious--an item can't be in two pigeonholes at the same time. This is a problem shared with hierarchical classifications; an item covering both addition and subtraction is still without a proper label. A second problem is that you have to construct the classification structure in advance. To use the postal analogy, you have to build the pigeonholes before you begin to sort the letters. Some guesswork is involved in choosing the number of pigeonholes needed, making them of the right scope, and making certain all the items will have a suitable classification. While it is not impossible to add or delete pigeonholes after the structure has been built, it is awkward.

The third problem concerns the names assigned to the pigeonholes. The structure allows only as many identifiers as there are pigeonholes, and thus there is no provision for separately retrieving two different items which have enough similarity to be placed in the same pigeonhole.

The hierarchical structure keeps adding detail until dissimilar traits serve to separate any items with significant differences; pigeonhole classifications do not. These two items:

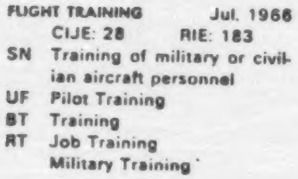
$$10 + 20 = ? \qquad \text{and} \qquad \begin{array}{r} 34 \\ +\ 55 \\ \hline ? \end{array}$$

are quite different. However, unless there was a pigeonhole for the first item named "Addition, whole numbers, no regrouping, 2 digit plus 2 digit, horizontal format, both ending digits zero, free response answer," it is likely that these items might end up in the same pigeonhole, even though they have important differences. Still, the pigeonhole approach has enough merit to warrant further consideration.

d. *Thesaurus Classification*. The dictionary defines a thesaurus as "a list of subject headings or descriptors, usually with a cross-reference system for use in the organization of a collection of documents for reference and retrieval." If one substitutes the word "items" for the word "documents" in the above definition, it appears that a thesaurus might be very useful for the structure of an educational test item collection.

Perhaps the most well-known thesaurus in education is the *Thesaurus of ERIC Descriptors*. ERIC, the Educational Resources Information Center, attempts to maintain a collection of most of the educational research and development information disseminated in the country. The collection is organized around and its documents retrieved by a list of authoritative descriptors, descriptors listed in the *Thesaurus*. Only by using words in the *Thesaurus* can a document be retrieved.

A simple example would be in order at this point. Suppose it would be helpful to know what information on flight training is already available. Turning to flight training in the Thesaurus, one finds this entry:



FLIGHT TRAINING Jul. 1966
CIJE: 28 RIE: 183
SN Training of military or civilian aircraft personnel
UF Pilot Training
BT Training
RT Job Training
Military Training

The scope note (SN) defines what ERIC means when it refers to the term "flight training." The used for (UF) note points out that "flight training" serves as the descriptor for "pilot training"; this note tells one that nothing in ERIC could be retrieved using "pilot training," since it is not a descriptor. The broader term (BT) note tells one that "flight training" is a subset of "training"; the implication is that if searchers find "flight training" to be overly narrow, they should proceed to "training." (ERIC also presents narrower term (NT) notes. For example, one finds, under "foreign policy" the note--NT: Imperialism.) Finally, the entry lists two related terms (RT)--"job training" and "military training." These are descriptors that may also interest someone searching for information on "flight training."

This type of classification has several advantages when applied to an item collection. First, it allows an item to be given more than one descriptor. Using ERIC, if the interest were in using tape recorders to teach foreign language, the searcher could cross "language instruction" with "audio equipment". Note that using either one of the descriptors alone would have given much irrelevant material, but that by using both, only the useful material under the intersection of the two groups is given. In an item bank, the two part addition and subtraction word problem that posed problems in the hierarchical and pigeonhole classifications could simply be assigned two descriptors, "whole number addition" and "whole number subtraction".

A second advantage is that descriptors can be added when the need arises. In both the hierarchical and pigeonhole structures, the classifications are set up in advance. A problem with collecting items for inclusion in any bank is that the content of the items generally isn't known until the items are collected. It is thus difficult to establish a fixed classification in advance. The thesaurus classification allows each item to be assigned as many descriptors as needed to describe the item. This lack of preconceived structure provides a flexibility which is very valuable. An additional benefit accruing from the lack of preconceived structure is that persons writing items for a hierarchical or pigeonhole system often write items to fill the classifications rather than writing the items they need. With the thesaurus system, item writers aren't guided into certain narrow areas.

Use of the thesaurus structure allows new descriptors to be established whenever the existing descriptors are inadequate. This does not mean that descriptors would be added haphazardly; every new descriptor proposed would be checked to be certain it is not simply a synonym for an existing descriptor. (For example, "subtraction with regrouping" eliminates the need for "subtraction with borrowing," or vice versa.) But new descriptors are easily added.

The thesaurus system does have a major disadvantage, however. If all one has to go on are several hundred descriptors, it is difficult to get a feeling for what is and is not in the data base. If trigonometry items are not in the collection, it would be better to know that in advance, rather than deducing it from unsuccessful searches for descriptors like "sine" and "tangent". But users of a thesaurus classification system often have only the list, usually in alphabetical order, of the descriptors to use.

Notwithstanding this problem, however, the thesaurus structure seems a promising way to classify incoming items.

Classifying Items for a Collection

a. Structure of the Item Collection. One conclusion which could be drawn from the previous section is that both the pigeonhole and thesaurus structures have advantages, and that they are better starting points than the hierarchical structure or no structure at all. Perhaps what is needed is a hybrid of the pigeonhole and thesaurus structures. Using a pigeonhole classification at a broad level could mitigate the thesaurus structure's failure to provide an overview of the collection's contents. Likewise, the thesaurus structure would add flexibility lacking with pigeonhole structures.

The NWREL structure uses "pigeonholes" to define the broad content of the collection. The collection has three content areas--reading, language arts, and mathematics. Each area is split into clusters--pigeonholes designed to help present a picture of what is included in the collection. The clusters have been reviewed by numerous educators and are now believed to be a comprehensive list of the content to be included in the collection. It should be possible to assign each appropriate item collected to one of the clusters. (There may still be some problem with multiple part items that bridge clusters, but the problem is much less severe at this broad level than it would be at a level of greater specificity. To date there has been little difficulty in determining how to assign items appropriate for multiple clusters. The list of clusters used in the NWREL collection is as follows:

I. READING

- A. Readiness
- B. Phonics
- C. Word Structure
- D. Vocabulary
- E. Comprehension and Analysis
- F. Study and Research Skills
- G. Oral Reading

II. LANGUAGE ARTS

- A. Capitalization
- B. Punctuation
- C. Grammar and Usage
- D. Sentence Structure
- E. Spelling
- F. Written Format
- G. Writing Sample
- H. Speaking Skills
- I. Listening Skills

III. MATHEMATICS

- A. Size and Characteristics
- B. Numeration
- C. Numerical Vocabulary
- D. Number Properties
- E. Whole Number Computation
- F. Fractions and Mixed Numbers
- G. Decimal Fractions
- H. Ratio and Proportion
- I. Percent and Percentages
- J. Integers
- K. Solving Equations and Inequalities
- L. Problem Solving Steps
- M. Measurement
- N. Geometry
- O. Geometric Quantities and Computations
- P. Roots and Exponents
- Q. Graphic Representation
- R. Probability
- S. Logic
- T. Statistics
- U. Mental Arithmetic
- V. Calculator Arithmetic
- W. Multiple Cluster Word Problems

No attempt was made to make the clusters parallel across areas (certainly reading comprehension is a "bigger" cluster than decimal fractions.) Nonetheless, the clusters should give a good idea of what content is contained in the bank. For instance, one can quickly tell that trigonometry will not be in the collection, but that geometry will be.

But beyond this level of specificity the pigeonhole structure loses its advantages. The lack of flexibility and difficulties with items fitting multiple pigeonholes become important. Therefore, from this level on, the thesaurus structure is used. As they are received, items are assigned one or more descriptors. Enough specificity is provided so the user can retrieve items in an efficient fashion.

The retrieval process begins with the user selecting one of the three areas, and then selecting the most appropriate cluster within that area. Once the cluster is chosen, the descriptors of the items in that cluster are presented, for example on a computer terminal. The user then chooses the most appropriate descriptors and the items having those descriptors are retrieved. Thus, the classification system uses a pigeonhole approach to bring the user into the relevant area and cluster, but a thesaurus approach in obtaining items having an appropriate descriptor.

b. Classification Terms. In real life, the classification of items will not be quite as straightforward as was described above. While use of the descriptors gives a good idea of the content of items, there is still a problem in identifying the format and structure of an individual item.

To use an earlier example:

$$10 + 20 = ? \qquad \text{and} \qquad \begin{array}{r} 34 \\ +\ 55 \\ \hline ? \end{array}$$

are both problems in whole number addition without regrouping. Both would probably be assigned the same descriptor. After all, the content of the items is whole number addition without regrouping; to answer the problems, a student needs to be able to add whole numbers. The differences are in the particular items chosen from the universe of whole number addition items. The point here is that descriptors relate to skills and content, not to the construction of individual items. This is not always a clear distinction, but it is an important one. It clearly points out the need for additional classification terms.

Accepted practice in objective referenced test construction also requires additional information. Educational measurement experts such as Popham, Hambleton, and others have spoken strongly of the need for a test item to be developed from a detailed set of test specifications.* Indeed, one can preduct that a major, and frequently justifiable, criticism which will be leveled against tests constructed from item banks is that the items are not derived from a comprehensive set of item specifications. This paper is not a suitable place to argue the merit of devoting large amounts of time to item specification development.

---

*See, for example, Chapter 6 of Popham, W. James. Criterion Referenced Measurement. Englewood Cliffs, NJ: Prentice-Hall, 1978.

Suffice it to say that writing an "acceptable" set of item specifications is a very time consuming task, that the cost-effectiveness of this step in terms of quality of the end product is not universally acknowledged, and that empirical research can and should be done to resolve this issue.

Regardless, the item bank concept decidedly is not in conflict with the desire for increased item specificity. While many existing item banks have poor classification schemes, there are also item bank structures that contribute greater specificity than could realistically be expected in the item specifications of most novice test developers. (Indeed, one of the most popular sources for the skill statements used by districts in their objectives and item specifications is, perhaps ironically, the content listings of commercially available item banks.)

In the next few paragraphs, the concept of limiters, item attributes, and response formats will be discussed. This assignment information, when combined with cluster and descriptor information, should provide the potential item bank user with an acceptably terse version of the information one would expect from an item specification. So while it is uncertain whether tests can be constructed more effectively starting from items or from item specifications, an item bank with adequate assignment information certainly allows either approach.

Shown on the next page is a reproduction of a card similar to the one used by the NWREL staff classifying the items. Each item has a completed card associated with it. Refer to the numbered areas of the sample card when reading the following descriptions.

ITEM BANK ASSIGNMENT CARD

Northwest Regional
Educational Laboratory

1 ☐ Receiver 2 ☐ Assigner 3 ☐ Reviewer

Area 4 ☐ Reading ☐ Language Arts ☐ Mathematics

Cluster 5 ______________________________ Aquisition Number 6 ☐

Descriptor(s) 7 ______________________________

8 ☐ Check here if descriptor approval is needed ______________________________

9 Limiter(s) ______________________________

____

10 Item Attribute(s) ______________________________

____

11 Response Format ☐ multiple choice ☐ matching ☐ free response/ single answer ☐ essay ☐ other ____________

☐ true/false ☐ fill-in ☐ free response/ short answer ☐ performance item ____________

12 Source Code ☐ A ☐ B ☐ C ☐ D ☐ E ☐ F ☐ ____________

13
Passage No. ☐ Graphic No. ☐ Statistics No. ☐
Problem No. ☐ Direction No. ☐ Other ☐ / ☐

1. Receiver - The initials of the person at NWREL who first processed the item as part of the item collection.

2. Assigner - The initials of the person completing the card; this person assigns descriptors and other terms to the item.

3. Reviewer - The initials of the person reviewing and accepting the technical adequacy of the item.

4. Area - Previously described.

5. Cluster - Previously described.

6. Acquisition Number - A unique number given to the item for reference and indexing purposes. The number would be between 00001 and 99999, and would not have any significance save as a reference for the item.

7. Descriptor(s) - Previously described.

8. Descriptor Approval - Assigners have to receive approval for any descriptor that they wish to use for the first time. Approval would be given only if the descriptor identified skills and content not already included in an existing descriptor.

9. Limiter(s) - Limiters are terms which are subsets of the content described by the descriptor. The differences in content between items with the same descriptors are described using the limiters. The assigner provides as many terms as seem appropriate to describe the item's content. Unlike descriptors, limiters are not strictly monitored (that is, a prescribed set of words is not used), and it is not possible to retrieve items based solely on the limiters. It is possible, however, to use the limiters to retrieve a subset of the items within a given descriptor. Examples of limiters for an addition problem might be the number of digits in the addends, whether one or more of the numbers include zeros and so on. It is important to remember that the assigner uses whatever limiters usefully describe the item and that no rigorous scrutiny of the limiter assignments are made. Limiters often convey the difficulty of the item; the readability of a passage or the complexity of a math problem would often be mentioned as a limiter.

10. Item Attribute(s) - The assigner uses the item attribute section to describe characteristics of the item not related to its content. The principal use is to list item format characteristics. For example, the use of functional transfer items for reading passages would be noted here; in mathematics, "horizontal format, decimals not aligned" might be two item attributes assigned to a decimal addition problem. Note that the item's content doesn't vary simply because it is presented in horizontal rather than vertical format. Content is specified by descriptors and limiters, item characteristics by item attributes. The attribute section is also used to point out oral items or items with prerequisite items.

11. Response Format - The means of answering an item would correctly be considered an item attribute, but since every item has a response format, it is given a separate section on the item bank assignment card. Eight common formats are listed with a blank available for listing other types.

12. Source Code - It may be useful in some cases to know the source from which an item was collected. The source code boxes are marked to indicate an item's source. For example, an item written by a district's teachers might be marked as source code A, items from the California Sample Assessment Exercises Manual as code B, and so on.

13. Other Card Numbers - In many cases, having the item (meaning the individual question) available is not sufficient. While 10 + 20 = ? is self contained, the question "What did LeRoy do once he came home?" is meaningless without the associated reading passage. And since that same passage probably serves for several other items, it is not efficient to copy the passage into the collection for each individual item. Instead, the passage, word, problem, or graphic (picture, chart, map, etc.) is given its own acquisition number and card. The number of the associated card goes in the appropriate box on the item card. Item directions are also given numbers. The direction "Solve the problem and circle the correct answer" may be shared by several hundred math problems; all the items with this direction have the same number in the direction box. Finally, it is expected that item difficulties and other statistics may be accumulated by NWREL for internal use. The statistics number box will reference the appropriate set of statistics.

c. Examples of Assignment Cards. Two examples of completed item bank assignment cards are given on the next pages. It should be remembered that the procedures demonstrated were successful in an initial application, but are amenable to change should a specific set of items so warrant.

1. Solve this problem and write the correct answer in the box.

10 + 20 = ☐

ITEM BANK ASSIGNMENT CARD

Northwest Regional
Educational Laborator

RS Receiver BLA Assigner SG Reviewer

Area ☐ Reading ☐ Language Arts ☒ Mathematics

Cluster Whole Number Computation

Aquisition Number 00241

Descriptor(s) Addition w/o Regrouping

☐ Check here if descriptor approval is needed

Limiter(s) 2 digit plus 2 digit, terminal digits both zeros

Item Attribute(s) Horizontal Format

Response Format ☐ multiple choice ☐ matching ☒ free response/single answer ☐ essay ☐ other
☐ true/false ☐ fill-in ☐ free response/short answer ☐ performance item

Source Code ☒ A ☐ B ☐ C ☐ D ☐ E ☐ F ☐

Passage No.

Graphic No.

Statistics No. 61271

Problem No.

Direction No. 00040

Other /

This card shows that the item was assigned number 00241, was reviewed for technical adequacy and was obtained from source A. The directions will be number 00040 and the statistics reference number will be 61271.

2. (Problem - Jane had 12 puppets and Cindy had 3. Marc had twice as many puppets as Cindy. How many puppets did Marc have?)

Which one of these choices shows a way to solve this problem?

a. 12 - 3 =
b. 3 + 12 =
c. 3 x 2 =
d. (12 + 3) x 2 =

ITEM BANK ASSIGNMENT CARD

Northwest Regional Educational Laboratory

SG Receiver EpA Assigner EgB Reviewer

Area ☐ Reading ☐ Language Arts ☒ Mathematics

Cluster Problem Solving Steps

Aquisition Number 10133

Descriptor(s) Translating to Math Equation

☐ Check here if descriptor approval is needed

Limiter(s) Whole number addition and Multiplication, one-step problem

Item Attribute(s) Extraneous information included

Response Format
☒ multiple choice ☐ matching ☐ free response/single answer ☐ essay ☐ other
☐ true/false ☐ fill-in ☐ free response/short answer ☐ performance item

Source Code ☐ A ☒ B ☐ C ☐ D ☐ E ☐ F ☐

Passage No.
Graphic No.
Statistics No. 31824
Problem No. 12818
Direction No. 00218
Other /

This card shows that the item was assigned number 10133, while the problem was assigned number 12818. (If this is, and will remain, the only item associated with the problem, the problem would be made part of the item, rather than being given a separate problem card.) The directions are on card 00218, while the statistics will be on card 31824.

Summary

To date, NWREL item assigners have classified approximately 20,000 items using the system described in this paper. Items from nearly twenty different sources have been accumulated under a common scheme, and the scheme has provided a consistent means of retrieving the items. The system is suitable for computerization; NWREL has already completed programming for automated storage and retrieval. Every item fits within the established structure. The work has been completed with less difficulty than occurred with smaller efforts using other classification structures. It is suggested that the item bank classification system described above be given serious consideration by educators setting up large item banks, especially when the items to be included are not written to relate specifically to a single set of objectives.

Item banks are a definite boon to educational testing, particularly when used for diagnostic and competency testing. One must consider, however, whether there is a second basis for the proliferation of relatively small item banks. Item banks appear to proliferate because of the perceived need to have the classification of the bank match extant goals and objectives. Would we not be better off to have users matching their objectives to a uniform classification system which would then lead to a substantial pool of items? It would seem so.

If there is a need to have item collections available for state and local education agencies, that need should be met. In meeting that need, a modicum of common sense and a little devotion to making a broadly usable product will go far toward making future item banks a significant and effective part of educational testing. The extent to which developers of banks can agree on common methods of organizing the collections is a direct indication of how useful the collections will be for general use.

The classification system proposed in this paper is far from ideal. It is offered to anyone who needs it as an acceptable interim structure until a better scheme can be found. It represents at least a start in redirecting funds from continual item development to effective use of high quality, comprehensive item collections. Given the current reductions of educational financing and predictions of even more adverse circumstances, the attempt to get more for our test development dollars is a critical step we can no longer afford to ignore.

Appendix B

May 5, 1981

Dear:

A little over a year ago, you applied for a temporary position as Technical Assistant or Technical Specialist on an item banking project being conducted by the Assessment and Measurement Program. We're back at it again! This time we have a need for ten Technical Assistants (Level 6) who are specialists in elementary reading and ten who are specialists in elementary mathematics. The persons selected will help us develop a 4th grade reading and math test. The work will be very brief--from eight to twenty hours total. We will be able to pay $7.50 an hour. There will be one introductory evening meeting before school is out, but the rest of the work will be performed during the summer at a mutually agreed upon time.

Since we already had a highly qualified group of applicants from our last item bank hiring procedure, we thought we would give you first opportunity to apply for the current position. If you would like to be considered, please telephone me to let me know. If we haven't heard from you by May 11, we'll assume your're not interested. If you have any questions, don't hesitate to contact me at 248-6800, ext. 368 (work) or 297-6765 (home).

Sincerely,

Evelyn J. Brzezinski
Coordinator of Contracted Services
Assessment and Measurement Program

EJB:pp

May 22, 1981

Dear :

Thanks for your willingness to participate in our new item banking project. Enclosed you will find several forms which you need to complete:

1. A Notice of Personnel Action, which you need to sign in the space marked Employee's Signature (at bottom of form). If your social security number is not typed in at the top of the form, please fill that in as well.

2. An Employee Data Sheet, which you should fill out completely. Please print clearly or type.

3. A W-4 form, which you should complete and sign.

A business reply envelope is enclosed for your convenience. When you send these signed forms back to us, also please include a copy of your resume. THIS IS VERY IMPORTANT; WE CANNOT PROCESS YOUR APPOINTMENT WITHOUT A RESUME.

I may or may not have talked with you by telephone to let you know the status of the project you will be working on. In case I haven't reached you, the main thing I need to let you know right now is that I am no longer a regular fulltime employee at Northwest Regional Educational Laboratory. Mike Hiscox and I have resigned in order to go into business on our own. The Lab has hired me, however, to participate in the study on which you will be working. Most of your contact with the Lab, though, will be with another staff member--still to be specified. That person will be the one to contact you about an initial meeting; but because of this change in staffing, it will still be a little while before that meeting is held.

The change in staffing does not have any implications for your involvement in the study. You will still be working between 8-20 hours during the summer, at a time you and the other four members of your "team" decide is mutually agreeable.

Please return the three forms and your resume as soon as you can. If you have any questions, don't hesitate to contact my secretary at the Lab, Barbara Hejtmanek (248-6800, ext. 366). If the new project director hasn't been named yet, she may not be able to give you a complete answer--but at least she can make a note of it and get back to you when the question can be answered.

Once again, I appreciate the fact that you volunteered to work with us again. I look forward to seeing you this summer.

Sincerely,

Evelyn J. Brzezinski
Assessment and Measurement Program

EJB:bah
Attachments

1. USE STATED FACTS TO ANSWER QUESTION

I. General Description

The student will be given a passage to read. From a list of alternatives the student will select the response that correctly answers a question concerning a time mentioned in the passage.

II. Sample Item

(From Stanford Diagnostic Reading Test, Green Level, Form A, page 14)

Did you know that bananas don't grow on trees? Even though the banana plant looks like a tree, it has no trunk or limbs. Bananas are grown in South and Central America and in India where the climate is hot and damp. They grow from flowers on the plant and are cut down before they get ripe. The banana plant dies down after just one crop. Since the banana does not have seeds like an apple or pear, new plants are grown from the roots of the old plants.

When are bananas picked?

a. Before they become flowers.
b. When they are ripe enough to eat.
c. When they form seeds.
*d. Before they are ripe.

Panel Comments

1. This paragraph is too difficult because it contains too much information.
2. The question here is an example of what to avoid in this lower level task--the paragraph uses "cut down" and the question asks "picked."

1. USE STATED FACTS TO ANSWER QUESTION

I. General Description

The student will be given a passage to read. From a list of alternatives the student will select the response that correctly answers a question concerning an object in the passage.

II. Sample Item

(From Belmont Standards Test System, Level C, Code 2304)

Do you live near the ocean? If so, you have probably eaten some type of seafood. One of the most popular seafoods is the clam. Maybe that's because clams are not only fun to eat, they're fun to dig for. Most kinds of seafood are taken out of the water, but you can dig for clams right on the beach. Clam diggers watch for little air holes in the sand that show the clam is hiding underneath, then turn a small shovelfull of sand over carefully, and presto! Fresh clams! Clams are good cooked all sorts of ways, but one of the best is over an open fire. Or maybe just the fun of being on the beach at sundown with a big fire blazing makes the clams taste better.

What does the writer say is one of the best ways to cook clams?

*a. Over an open fire.
b. In a big kettle.
c. In the oven.
d. The writer doesn't like clams.

Panel Comments

1. The panel did not like this paragraph. It is too emotional. It also seems to be talking down to the student.

2. The question is not a "what" question. It is more like a "how" question. A better question would be "What do clam diggers watch for?"

3. Distraction "d" was considered a poor distraction.

1. USE STATED FACTS TO ANSWER QUESTION

I. General Description

The student will be given a passage to read. From a list of alternatives the student will select the response that correctly answers a question concerning a location mentioned in the passage.

II. Sample Item

(From Belmont Standards Test System, Level C, Code 2304)

Do you live near the ocean? If so, you have probably eaten some type of seafood. One of the most popular seafoods is the clam. Maybe that's because clams are not only fun to eat, they're fun to dig for. Most kinds of seafood are taken out of the water, but you can dig for clams right on the beach. Clam diggers watch for little air holes in the sand that show the clam is hiding underneath, then turn a small shovelfull of sand over carefully, and presto! Freah clams! Clams are good cooked all sorts of ways, but one of the best is over an open fire. Or maybe just the fun of being on the beach at sundown with a big fire blazing makes the clams taste better.

Where do you look for clams?

a. In the ocean.
b. Wherever there is sand.
*c. Right on the beach.
d. Under big rocks.

Panel Comments

1. The panel did not like the paragraph. A more factual paragraph is needed.
2. The panel discussed here whether the instructions to the items would be "Choose the best answer." It was decided that for the test, the best response would be called for. Therefore, a distraction could be close but still not as good as the correct response.
3. Distraction "d" is a good distractor. It is not mentioned in the paragraph, but is a plausible response if the student has not read the passage.

2. DETERMINE SEQUENCE OF EVENTS IN A PASSAGE

I. General Description

The student will be given a passage to read. From a list of alternatives the student will select the response that identifies something that occurred before a stated event.

II. Sample Item

(From Belmont Standards Test System, Level C, Code 2312)

One day when Sam and Hank were playing kickball, Sam fainted. Hank ran to get his mother, and she knew just what to do. She laid him down flat on his back, and gently turned his head to one side. Then she opened his shirt at the collar so he could breathe more easily. She held something strong under Sam's nose. It made Hank's eyes water. Later, she told Hank it was ammonia. As soon as Sam felt better, Hank's mother made lemonade. Then things didn't seem so bad.

What happened before Sam fainted?

a. He laid down on his back.
*b. He was playing kickball.
c. His mother made lemonade.
d. He smelled some ammonia.

Panel Comments

1. The topic of this paragraph was considered good for fourth graders.

2. "Before" items also include "What happened first?" or "What happened yesterday?".

3. The panel also discussed that one item type not presented here in the item specifications was one where the student was given five events from the passage and asked to put them in order:

1. ________
2. ________
3. ________
4. ________
5. ________

Which of these sets of numbers shows the correct order of events as they appeared in the paragraph?

a. 4-2-1-3-5
b. 1-2-3-4-5
c. 2-4-3-1-5
d. 3-5-1-4-2

The panel wanted two items of this type: one passage that contained instructions or directions and one that had events like this paragraph on Sam and Hank.

2. DETERMINE SEQUENCE OF EVENTS IN A PASSAGE

I. General Description

The student will be given a passage to read. From a list of alternatives the student will select the response that identifies something that occurred after a stated event.

II. Sample Item

(From Belmont Standards Test System, Level C, Code 2312)

One day when Sam and Hank were playing kickball, Sam fainted. Hank ran to get his mother, and she knew just what to do. She laid him down flat on his back, and gently turned his head to one side. Then she opened his shirt at the collar so he could breathe more easily. She held something strong under Sam's nose. It made Hank's eyes water. Later, she told Hank it was ammonia. As soon as Sam felt better, Hank's mother made lemonade. Then things didn't seem so bad.

What happened after his mother held the ammonia under Sam's nose?

a. Sam fainted.
b. Sam's mother opened his shirt collar.
*c. Hank's eyes watered.
d. Hank played kickball.

Panel Comments

1. "After" questions can also include "What happened last?" or "What will happen tomorrow?".

3. IDENTIFY CAUSE AND EFFECT RELATIONSHIPS

I. General Description

The student will be given a passage to read. From a list of alternatives the student will select the response that identifies the cause of an event that occurs in the passage. The cause will be mentioned in the passage.

II. Sample Item

(From STEP, Level F, Form X, page 6)

The Fulani live in Africa. They travel in small family groups, hardly ever staying in one place for more than two or three weeks. It is the task of the head of the group to decide where the people will move their herds for fresh pasture and water. When the group reaches a new grazing place, the women and girls quickly weave small grass houses which the families sleep in. When they move on, the houses are left behind and new ones are made in the next camping place.

The Fulani move from place to place because they

a. are chased away by their enemies.
*b. need food for their animals.
c. enjoy traveling to new places.
d. have to keep looking for new jobs.

Panel Comments

1. The panel considered this an interesting paragraph.

2. The panel was concerned that the Fulani are never identified as a tribe of people living in Africa.

3. Of the two passages given in the item specifications, both questions ask for the cause given the effect. There needs to be one asking the effect given the cause. An example for this paragraph would be "What do the Fulani do because their animals need food?". Answer: "Move from place to place."

3. IDENTIFY CAUSE AND EFFECT RELATIONSHIPS

I. General Description

The student will be given a passage to read containing both a cause and its effect. The student will select from a list of alternatives the response that identifies the cause of an event that occurs in the passage.

II. Sample Item

(From Metropolitan Achievement Tests, Elementary Level, Form JI, page 13)

Nancy and Rick had three turtles. "I wonder which is the fastest?" asked Dale.

"Let's have a race and see," said Nancy.

Rick made two lines about three feet from each other on the sidewalk. Then he put the turtles on one line, and carefully pointed them toward the second line. But when the turtles moved, they crawled in different directions.

Then Nancy drew a large round circle and placed the turtles in the center. "Now," she said, "no matter which way they crawl, it will surely be toward the finish line."

Rick's idea didn't work because the

a. turtles didn't move.
b. lines were too close together
c. sidewalk was too cold.
*d. turtles crawled different ways.

Panel Comments

1. This is a cute passage. Students will like this passage better than the passage about the Fulani because this one is light and enjoyable.

4. DETERMINE OBJECT OF A REFERENT

I. General Description

The student will be given a passage to read. The student will be given a pronoun that is referenced in the passage. From a list of alternatives the student will select the response that correctly identifies the referent in the passage.

II. Sample Item

(From Belmont Standards Test System, Level C, Code 2316)

One day at school, Miss Blake told the class how important it was to eat right. "You are what you eat," she said. Ron thought about that a long time. That night at dinner, his mother served carrots. Ron didn't want to eat any, but he was afraid his mother would be angry. He ate four. When he went to bed he couldn't sleep. He kept feeling the top of his head, feeling for the green tops to sprout up. He looked at his legs to see if they were turning orange. After a while, he closed his eyes. Suddenly, he could feel it happening. . .slowly, his legs got longer, longer. . .his body got skinny and orange. . .his eyes and ears disappeared. . .the sprouts pushed out of his head. . .then he screamed, and sat straight up in bed, holding his hair in his hands. He ran to the mirror. . .it was morning, and he was back to normal.

In the seventh line of the passage, the word they is underlined. What does they refer to?

a. Carrots
b. Other students in Ron's class
c. The green tops
*d. Ron's legs

Panel Comments

1. Good passage. It is interesting for students.

2. This passage will also be good for the later task of determining fact and fiction.

3. Nicknames could also be used for a referent. For example, if a child is referred to as "Champ" after winning, then the question could ask who "Champ" refers to.

4. Brand names as referent had originally been discussed for referent, but it was decided that the students could identify the object (e.g., Tide, Kleenex, Nike) without reading the passage.

5. IDENTIFY MAIN IDEAS

I. General Description

The student will be given a passage to read. The student will select from a list of alternatives a title for the passage that best reflects the main idea.

II. Sample Item

(From Stanford Diagnostic Reading Test, Green Level, Form A, page 14)

Did you know that bananas don't grow on trees? Even though the banana plant looks like a tree, it has no trunk or limbs. Bananas are grown in South and Central America and in India where the climate is hot and damp. They grow from flowers on the plant and are cut down just after one crop. Since the banana does not have seeds like an apple or pear, new plants are grown from the roots of the old plants.

Which title best tells what this paragraph is about?

a. Bananas and Other Fruits
*b. How Bananas Grow
c. The Banana Farm
d. Who Eats Bananas?

Panel Comments

1. The type of item missing from the examples given is one that asks "Which sentence from this selection gives the main idea?" and the responses include the topic sentence of the paragraph.

5. IDENTIFY MAIN IDEAS

I. General Description

The student will be given a passage to read. From a list of alternatives the student will select the sentence that best reflects the main idea of the passage.

II. Sample Item

(From SRA Achievement Series, Level D, Form 1, page 10)

"Where's the cat?" dad asked. "We can't leave until we take him next door to stay with the Kaplans."

"Tink was on the porch," answered Jess.

"On the porch!" dad said, frowning. "Jess, I asked you to keep the porch door locked. That cat pushes against the unlocked door and swings it open. Then he disappears for the day."

"I did lock the door, dad," Jess said.

Mother shouted, "Don't blame Jess! Our suitcases were piled in the hall, and you carried them to the car. You left the porch door open between trips to the car."

"So I did," dad said, "and I'm sorry, Jess."

Suddenly, they heard a noise from the car. "I think I know where Tink is hiding," Jess said.

What is the main idea of this story?

a. Jess should learn to take better care of Tink.
b. Tink runs away because his family is leaving.
c. Tink hides in the car so that he can go along.
*d. Dad realizes it is his fault that Tink is missing.

Panel Comments

1. The nickname "Tink" would be an example of determining the object of a referent.

5. IDENTIFY MAIN IDEAS

I. General Description

The student will be given a passage to read. From a list of alternatives the student will select the word that best reflects the main idea of the passage.

II. Sample Item

(From Gates-MacGinitie Reading Tests, Level D, Form 3, page 4)

On top of the pine tree was a deserted hawk's nest. Every child in Toronto knew the nest, but no one had ever seen a sign of life about it. There it was, year after year, ragged and old. Yet, strange to tell, in all that time it never did fall to pieces, like other old nests.

This story is mainly about a

a. tree
*b. nest
c. child
d. hawk

Panel Comments

1. For distractors in this type of item make sure they all appear in the passage and the correct answer should disapportionately appear more than the incorrect responses.

2. There was concern about including the word "Toronto." Students may find its pronunciation difficult and that could distract from the comprehension of the paragraph.

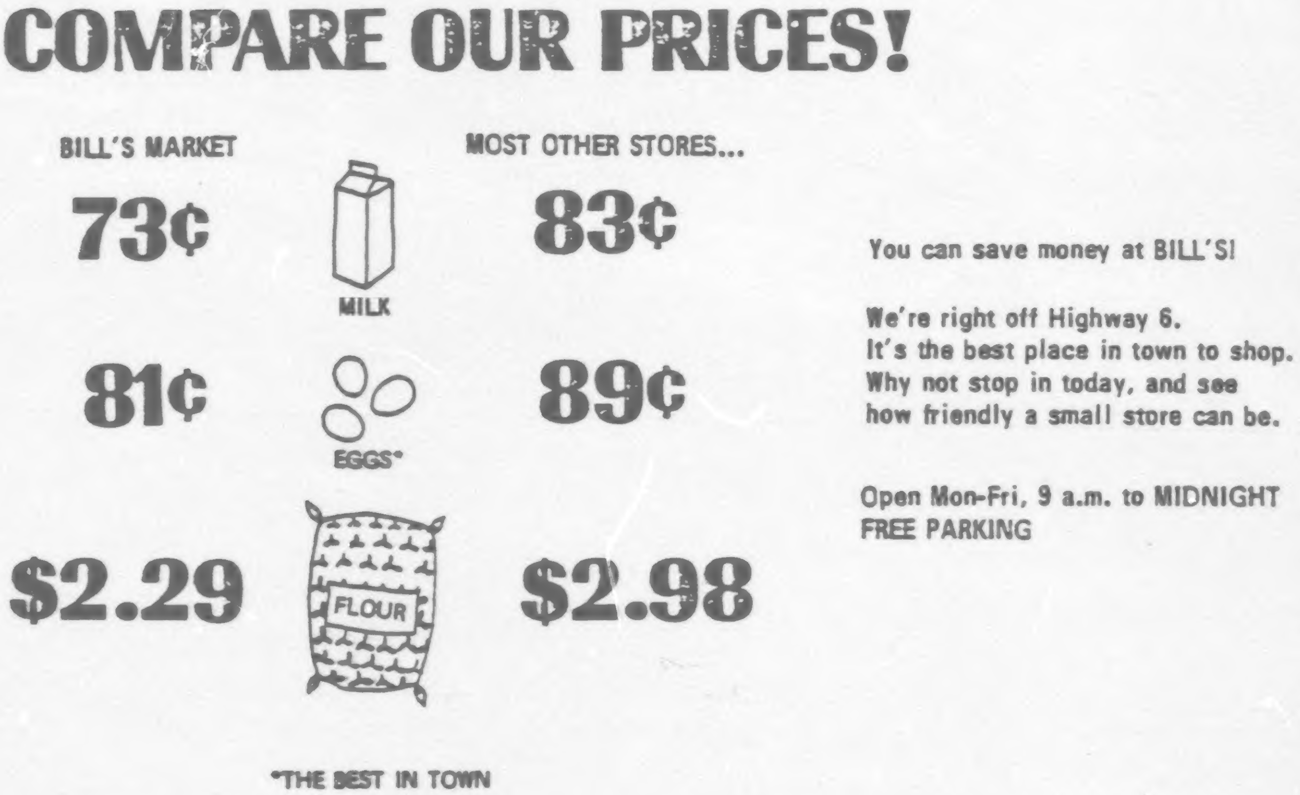
6. IDENTIFY CONCLUSIONS AND SUPPORTING FACTS

I. General Description

The student will be given a passage to read. The student will be asked to select from alternatives a conclusion that the facts in the passage support.

II. Sample Item

(From Belmont Standards Test System, Level C, Code 2324)



COMPARE OUR PRICES!

BILL'S MARKET		MOST OTHER STORES...
73¢	MILK	83¢
81¢	EGGS*	89¢
$2.29	FLOUR	$2.98

*THE BEST IN TOWN

You can save money at BILL'S!

We're right off Highway 6.
It's the best place in town to shop.
Why not stop in today, and see
how friendly a small store can be.

Open Mon-Fri, 9 a.m. to MIDNIGHT
FREE PARKING

The ad says Bill's is "the best place in town to shop." What might make you think this is true?

- a. Bill's is open Monday to Friday.
- *b. Prices seem lower at Bill's.
- c. Bill's is the only store with milk, eggs and flour.
- d. Bill's has free parking.

Panel Comments

1. The panel felt that this may be testing the student's graph reading skills. It would be better if the information were presented in paragraph form. An example that the panel liked can be found in the California Achievement Test, Level 14C, page R5.

6. IDENTIFY CONCLUSIONS AND SUPPORTING FACTS

I. General Description

The student will be given a passage to read. The student will be given a conclusion based on the passage and will be asked to select from alternatives the response that best represents facts from the passage that support that conclusion.

II. Sample Item

(From Belmont Standards Test System, Level D, Code 2324)

Around 4 p.m. the second Thursday in August, Margaret quite suddenly got the notion that it would be fun to stop by and see Dr. Baily, the counselor. Unfortunately, he was not in. She read the following sign on his door:

John Baily, Counselor
Family, Personal,
Marriage

Hours: 9 to 5 Mon., Tues., Wed.
9 to Noon Fri. and Sat.

Call for Appointment
292-7000

--Have a Happy Day--

What part of Dr. Baily's sign best supports Margaret's conclusion that Dr. Baily was friendly?

a. Call for Appointment
*b. Have a Happy Day
c. Family, Personal, Marriage
d. John Baily, Counselor

Panel Comments

1. This passage is not in the students' world of recognition. Students may not be familiar with the idea of counselors.
2. The item here involves inference.
3. Be careful to avoid making these cause and effect questions.
4. An example that the panel liked is from the Stanford Diagnostic Reading Test, Green Level, Form M, page 15, item number 35.
5. A detective story would be good here. Or why does Johnny's mother think he has chicken pox. Or two friends discussing why they like a third person (conclusion) because of certain characteristics (supporting facts). Or why a sports player is considered good.

7. DISTINGUISH FACT, OPINION, FICTION, NON-FICTION

I. General Description

The student will be given a passage to read. From a list of alternatives that include statements from the passage, the student will select the alternative that is not an opinion.

II. Sample Item

(From California Achievement Tests, Level 15C, page R9)

I am writing about the Pen and Ink Club's new art show. The show is titled "Creatures in Ink." It will be shown for three weeks. It's the best show Oliver School has ever had. Everyone who sees it will agree. There are ink drawings of dinosaurs, birds, lizards, spiders, and monkeys. These beautiful pictures now decorate some of the halls of our school. Many students are saying that the Pen and Ink Club should be given a special prize. The whole school loves "Creatures in Ink."

Andrew Warrel
Secretary, Pen and Ink Club

Which of the following is not an opinion?

*a. The art show is titled "Creatures in Ink."
b. The whole school loves "Creatures in Ink."
c. The club should be given a prize for "Creatures in Ink."
d. The "Creatures in Ink" pictures are beautiful.

Panel Comments

1. The panel liked using this format for passages.

2. The panel suggested replacing "not an opinion" with "a fact."

3. Other types of passages that could be used are ones involving advertisements or propaganda.

4. There was some discussion about the use of an earlier passage using a dream. The comment was made that putting a dream within nonfiction is confusing.

5. One way of the students identifying fiction is to ask "Which things in this story make it appear not real?"

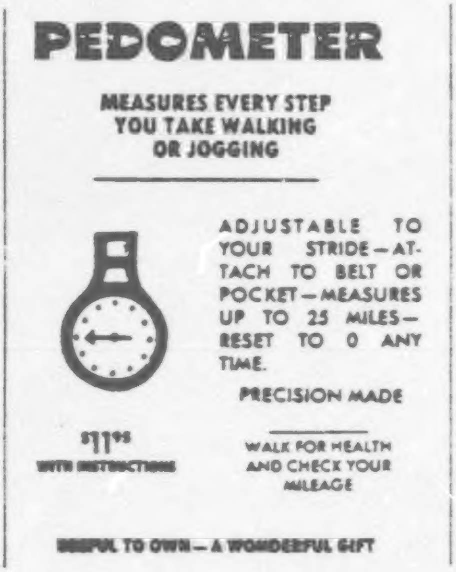
8. INFER FORMULATIONS (CONCLUSIONS, GENERALIZATIONS, UNSTATED ASSUMPTIONS)

I. General Description

The student will be given a passage to read. The student will select from a list of alternatives a conclusion that can be inferred from the passage, that is, the conclusion is not actually stated.

II. Sample Item

(From Belmont Standards Test System, Level D, Code 2332)



PEDOMETER

**MEASURES EVERY STEP
YOU TAKE WALKING
OR JOGGING**

ADJUSTABLE TO YOUR STRIDE—ATTACH TO BELT OR POCKET—MEASURES UP TO 25 MILES—RESET TO 0 ANY TIME.

PRECISION MADE

$\$11^{95}$
WITH INSTRUCTIONS

WALK FOR HEALTH
AND CHECK YOUR
MILEAGE

USEFUL TO OWN—A WONDERFUL GIFT

What happens to the pedometer after it measures 25 miles?

a. It has to be replaced.
*b. It has to be reset.
c. It wears out.
d. It sounds an alarm.

Panel Comments

1. An example of inferring an unstated assumption was found in the California Achievement Test, Level 14C, page R13, item number 60.

9. GENERATE SOLUTIONS, PREDICT OUTCOMES, APPLY IDEAS TO NEW SITUATIONS

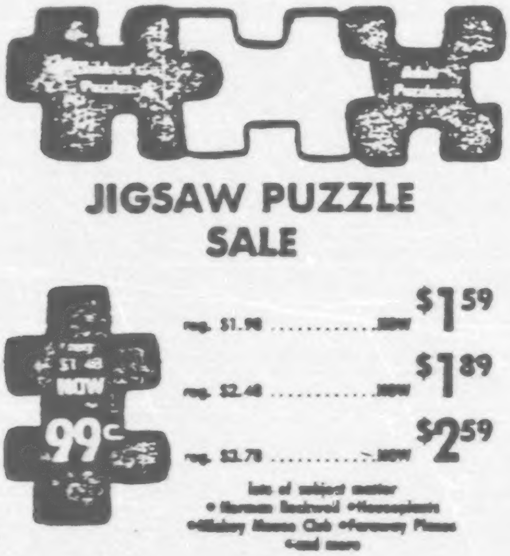
I. General Description

The student will be given a passage to read. The student will select from a list of alternatives the solution to a problem that has been presented in the passage.

II. Sample Item

(From Belmont Standards Test System, Level C, Code 2328)

Molly needed a present for her friend Fran, but she only had $3 to spend, and it was hard to think of what to buy. She knew Fran liked games, puzzles and books. But she couldn't find just the right gift. Then she saw this ad in the paper:


JIGSAW PUZZLE
SALE
NOW 99¢
reg. $1.98 ....NOW $1.59
reg. $2.48 ....NOW $1.89
reg. $3.78 ....NOW $2.59

What do you think Molly will do?

- *a. Buy Fran a puzzle.
- b. Try to earn more money for a present.
- c. Tell Fran about the sale.
- d. Tell Fran she can't afford a present.

Panel Comments

1. This would be a better problem if there had been a toy airplane for $1.89, a puzzle for $2.59 and a game for $3.49.

9. GENERATE SOLUTIONS, PREDICT OUTCOMES, APPLY IDEAS TO NEW SITUATIONS

I. General Description

The student will be given a passage to read. From a list of alternatives the student will select what will most likely happen next in the passage.

II. Sample Item

(From SRA Achievement Series, Level D, Form 1, page 8)

Kelly tried to paint in his bedroom. His mother said, "You can't paint here."

In the living room his brother yelled , "I'm watching TV. You're in my way!"

In the kitchen his sister said, "You can't paint here! I'm doing my homework."

Kelly went outside and sat down beside the doghouse.

"I need a place to paint," Kelly told the dog.

He put his paint box down and got out his paper. Then he dipped his brush in the dog's water dish.

What will probably happen next in this story?

*a. Kelly will start painting.
b. Kelly will do his homework.
c. Kelly will decide not to paint.
d. Kelly will go back to the kitchen.

9. GENERATE SOLUTIONS, PREDICT OUTCOMES, APPLY IDEAS TO NEW SITUATIONS

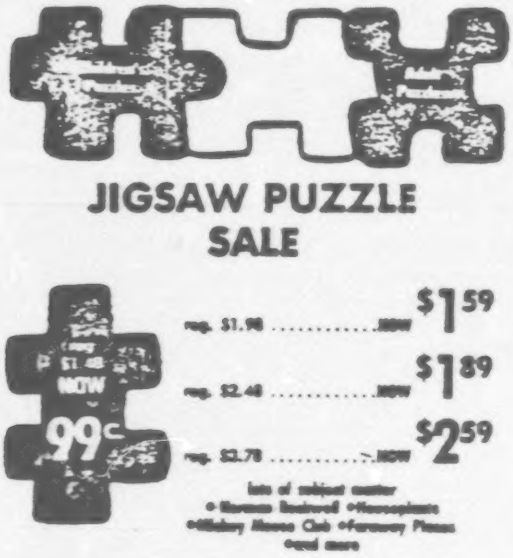
I. General Description

The student will be given a passage to read. The student will select from a list of alternatives the action that will be taken next time this situation occurs.

II. Sample Item

(From Belmont Standards Test System, Level C, Code 2328)

Molly needed a present for her friend Fran, but she only had $3 to spend, and it was hard to think of what to buy. She knew Fran liked games, puzzles, and books. But she couldn't find just the right gift. Then she saw this ad in the paper:


JIGSAW PUZZLE
SALE
NOW
99¢
reg. $1.98 .... NOW $1.59
reg. $2.48 .... NOW $1.89
reg. $2.78 .... NOW $2.59

Next time Molly needs an idea for a gift, what do you think she might do?

- a. Ask Fran for an idea.
- b. Save up until she has lots of money.
- *c. Look in the paper.
- d. Shop until she finds something.

Reading Test

DIRECTIONS

Read each story and then answer the items following it. Choose the best answer for each item. Then mark the space on your answer sheet with the same letter as the answer you select.

SAMPLE STORY

Susan was very sleepy. The family had been riding in the car for over five hours. She would be very glad when they got home. Then she could sleep in her own bed.

SAMPLE ITEM

Where was Susan?

A. In bed
B. In a bus
C. In a car
D. In an office

You should have filled in the space "C" on your answer sheet because Susan was in a car. Your answer sheet would look like this:

⊏A⊐ ⊏B⊐ ■ ⊏D⊐

Use a #2 pencil to mark your answers.

If you make any changes, be sure to erase completely.

IS

Buy Piggaloo!

Your kids will play it for hours and hours. Make your children better spellers. Two to four can play. For children ages 7 to 12.

---

1. Which of the following is a fact?
   - A. Many people have bought Piggaloo.
   - B. Children play it for hours and hours.
   - C. Children who play it are better spellers.
   - D. Two to four people can play

2. In the first sentence what does "it" mean?
   - A. Piggaloo
   - B. Children
   - C. Hours
   - D. Kids

3. What can parents expect when they buy Piggaloo?
   - A. Teenagers will like Piggaloo.
   - B. It takes hours and hours to learn to play.
   - C. Their children will become better spellers.
   - D. Three children cannot play.

Ellen knew that Jan's birthday was coming soon. Ellen had saved $4.00 to buy her a present. She knew Jan liked puzzles and model airplanes, but she could not find just the right gift. One day Ellen's mom showed her an ad in the paper.

CLOSE-OUT SPECIAL

Books	$2.98
Puzzles	$3.50
Models	$5.00 and up

---

4. In the second sentence, "her" refers to

   A. Jan.
   B. Ellen.
   C. Ellen's mom.
   D. the newspaper ad.

5. What will Ellen probably do for Jan's present?

   A. Buy a book
   B. Buy a puzzle
   C. Buy a model
   D. Give Jan $4.00

6. Which of the following sentences best describes the story?

   A. Jan has a birthday.
   B. Ellen's mom finds an ad.
   C. Ellen looks for a birthday present.
   D. Ellen saves $4.00

7. What did Ellen's mom show Ellen?

   A. A book
   B. A puzzle
   C. A model airplane
   D. An ad

The common use of airplanes for travel is fairly new. Airplanes were developed in the early 1900's. However, the planes were small and flying was uncomfortable. Lack of radios made navigation difficult and people were uncertain of their safety. Airplanes were used mainly as fighters in the two World Wars.

In the 1950's larger, more comfortable planes with jet engines were developed. Traveling by plane became common. Today flying is as exciting as it is safe.

---

8. The story states that we now have large comfortable jets. What result does this have?

   A. Airplanes were developed.
   B. The author would like to fly in an old airplane.
   C. People buy airplanes.
   D. Flying is common.

9. Which statement is an opinion?

   A. Planes were used as fighters.
   B. Now we have large jets.
   C. Very few people used planes for travel in the early 1900's.
   D. Today flying is exciting.

10. In the early days of flying, planes were small and uncomfortable. Navigation was difficult. People worried about safety. What effect did this have?

    A. Planes were used as fighters.
    B. Jets were developed.
    C. Few people traveled by plane.
    D. It would be fun to fly in an old plane.

They landed with with a "Thump!" Izzy looked at Fizzy. Fizzy slithered off her bed to a small hole in the bottom of the ship. She opened the cover over the hole, crawled down the side of the spaceship and promptly disappeared behind a blade of grass.

"Come quickly, Izzy," she called. "I have never seen this color before! It's like blue, but more yellow. The entire plant is all one color. It doesn't have any leaves, just one tall, slender trunk."

Izzy looked down from the spaceship. "I have studied the Earth," he said. "We are in a huge forest of grass."

---

11. Look at the following statements. Then indicate the order in which they occurred according to the story.

    1. Fizzy crawled down the side of the spaceship.
    2. Izzy looked at Fizzy.
    3. The spaceship landed.
    4. Fizzy slithered to a hole in the ship.

    A. 3, 4, 1, 2
    B. 2, 1, 4, 3
    C. 1, 2, 4, 3
    D. 3, 2, 4, 1

12. Which animal is Fizzy closest to in size?

    A. An elephant
    B. A cat
    C. A horse
    D. An ant

13. What will Izzy probably do next?

    A. Disappear behind a blade of grass
    B. Crawl down the spaceship
    C. Study the earth
    D. Take off in the spaceship

14. In the second paragraph what is Fizzy describing?

    A. A tree
    B. Grass
    C. Izzy
    D. The color blue

15. What part of the story best supports the idea that Izzy and Fizzy are from outer space?

    A. Fizzy slithered off her bed.
    B. Izzy and Fizzy looked at one another.
    C. Fizzy crawled down the side of the spaceship.
    D. She opened the cover over the hole.

16. Which of the following statements tells you the story is fiction?

    A. They landed with a "Thump!"
    B. Izzy looked at Fizzy.
    C. The entire plant is all one color.
    D. We are in a huge forest of grass.

Father's Day was approaching. Emily and Tom wanted to surprise their dad. They did not have any money for a present, though. "Let's make Dad some brownies," said Emily. "He really likes chocolate, and Mom rarely makes dessert."

"Brownies taste good," said Tom. "I like them a lot."

Tom found a cookbook, and they looked up a recipe. Emily gathered the ingredients. Tom got out bowls, pans, and a mixer. Soon the brownies were mixed and in the oven. Tom set the timer. After 10 minutes, Emily said, "Oh no, I forgot baking soda. That's what makes the brownies rise!"

---

17. Who found a cookbook?

    A. Mom
    B. Dad
    C. Tom
    D. Emily

18. Why did Tom and Emily bake brownies?

    A. They were hungry.
    B. They were bored.
    C. They liked to cook.
    D. They wanted to surprise their Dad.

19. Why did they think Dad would like brownies?

    A. Tom liked brownies.
    B. Father's Day was approaching.
    C. Dad liked chocolate.
    D. Tom found a cookbook.

20. Which statement is an opinion?

    A. Father's Day was approaching.
    B. Brownies taste good.
    C. Tom found a cookbook.
    D. Oh, no, I forgot baking soda.

21. The best title for this passage is

    A. A Father's Day Present
    B. Tom Likes Brownies
    C. How to Make Brownies
    D. Mom Makes a Dessert

22. Tom and Ellen are

    A. friends
    B. parents
    C. cousins
    D. brother and sister

23. The next time they bake brownies, what should Tom and Ellen do differently?

    A. Make more brownies
    B. Double check the ingredients
    C. Save money for a present
    D. Have Mom make more desserts

24. What happened before Emily found the ingredients?

    A. Tom found a cookbook.
    B. They mixed the brownies.
    C. Tom set the timer.
    D. Tom got out bowls.

On the third day of our vacation, B.D. and I met Jerry. He was standing in a creek and vacuuming rocks, gravel, and sand from the bottom. Then he walked up on the bank and showed us how to pan for gold. First he filled a metal pan with some of the materials he had vacuumed from the creek bottom. Then he scooped in some water. "Slosh it around slowly to let any gold settle," he said. "Then slowly spill out the water and rocks."

I carefully followed his instructions. After ten minutes my pan was almost empty. That's when I saw three small flakes of gold. I was so excited I almost dropped the pan. "Let's come here next year for our vacation," I said to B.D.

---

25. When panning for gold, what order do you follow?

    1. Slosh water in the pan slowly.
    2. Spill out rocks and water.
    3. Vacuum the river bottom.
    4. Put material from the river bottom in a metal pan.

    A. 3,1,4,2
    B. 2,4,3,1
    C. 3,4,1,2
    D. 4,1,2,3

26. Why did Jerry slosh water around slowly?

    A. So he wouldn't get wet
    B. To let any gold settle
    C. So he wouldn't kill any fish
    D. To keep gravel from falling out

27. Which sentence best summarizes this story?

    A. It is important to follow instructions.
    B. You can make money on your vacation.
    C. I learned to pan for gold.
    D. I spent my vacation panning for gold.

28. "B.D." refers to

    A. the author's dog.
    B. the author.
    C. Jerry's friend.
    D. the author's friend.

29. If you wanted to pan for gold, you would go to

    A. the desert.
    B. an abandoned gold mine.
    C. a river with a muddy bottom.
    D. a river with a rocky bottom.

Blue skies turned gray. Clouds rolled in. The birds hushed. The hair on Chris's arm stood on end as static electricity built up. The wind whipped the lake into towering waves. It hurled the sand into Chris. "Hurry and get in the car! Wrap your towel around you! Hurry!" someone called. Chris ran to the car, but the sand kept stinging. Tears came to Chris's eyes.

---

30. What happened before the birds hushed?

    A. Clouds rolled in.
    B. The wind started to blow.
    C. Chris ran to the car.
    D. Tears came to Chris's eyes.

31. What happened because of the static electricity?

    A. Blue skies turned gray.
    B. The hair on Chris's arm stood on end.
    C. The wind started to blow.
    D. Chris ran to the car.

32. How did Chris feel at the end of the story?

    A. Happy
    B. Friendly
    C. Curious
    D. Scared

33. What do you think will happen next?

    A. The sky will turn blue.
    B. There will be a rainstorm.
    C. Birds will start to sing.
    D. There will be an earthquake.

Loca noticed that the water was getting colder and colder. Like all whales, he was protected from the cold by a layer of fat. But the cold water killed the krill, tiny animals that gray whales eat. "It's time to move south," he thought. He slapped his tail against the water five times. "Let's go" he signaled. The other whales followed him.

As he started to swim, Loca sang a song. It went:

"Sun is low,
Water is cold,
Let's go south to warm water."

Twice a year gray whales like Loca migrate over 3,000 miles. They spend the summer near Alaska where there are many krill to eat. In the winter they swim south to Mexico or Hawaii. The warm water there also contains krill.

The sounds that whales make are called songs. Many people enjoy listening to these songs. Different kinds of whales sing different songs. Gray whale songs are interesting to hear.

---

34. This article is about

    A. songs.
    B. krill.
    C. whales.
    D. swimming.

35. The water was getting colder. Krill were being killed. What did this make the whales do?

    A. Sing
    B. Move north
    C. Move south
    D. Eat krill

36. Which of the following statements is an opinion?

    A. Twice each year gray whales migrate.
    B. Whales spend the summer near Alaska.
    C. People who study whales have heard them sing.
    D. Gray whale songs are interesting to hear.

37. Why do the whales migrate?

    A. So they can sing
    B. To find krill
    C. Because the sun is low
    D. They become cold

38. Which of the following happened after Loca slapped his tail against the water?

    A. The water got colder.
    B. The krill died.
    C. Loca sang a song.
    D. Loca swam north.

39. Which paragraphs of the story contain fiction?

    A. 1 and 2
    B. 1 and 3
    C. 2 and 3
    D. 3 and 4

"Going to the mountains is going home." This was written by John Muir. Muir was a quiet man who loved the mountains. Much of his life was spent trying to preserve the wilderness.

Muir was born in Scotland. In 1868 he moved to California where he lived the rest of his life. Muir loved to spend weeks in the California mountains. He would carry only his notebook, fried bread, and tea.

Muir felt that people must have wilderness areas. He convinced many people there was a need. His writing about the wilderness resulted in the establishment of three national parks.

Muir died in 1914, but he is still remembered today. The Sierra Club, a group he founded in 1892, is still devoted to saving the wilderness. Many people continue to read the books he wrote. After his death, a trail was developed and named for him, the John Muir Trail.

---

40. Where was John Muir born?

    A. Scotland
    B. The mountains
    C. California
    D. Sierra

41. What caused the establishment of three national parks?

    A. Muir started the Sierra Club.
    B. A John Muir Trail was created.
    C. Muir moved to California.
    D. Muir wrote about the wilderness.

42. When was the Sierra Club started?

    A. 1868
    B. 1892
    C. 1914
    D. 1920

100

43. How do we know John Muir loved the wilderness?

    A. He moved to California.
    B. He was a quiet man.
    C. People still read his books.
    D. He spent weeks in the mountains.

44. Which happened last?

    A. The Sierra Club started.
    B. Muir moved to California.
    C. Muir died.
    D. The John Muir Trail was developed.

45. The best title for this story is

    A. The Life of John Muir
    B. Three National Parks
    C. Born in Scotland
    D. Fried Bread and Tea

Appendix F

1. ADD WHOLE NUMBERS WITHOUT REGROUPING

I. General Description

Given an addition problem requiring no regrouping and consisting of a 3-digit number and a 1-digit number presented in a vertical format, the student will select the correct answer from a list of alternatives.

II. Sample Item

$$\begin{array}{r} 395 \\ +\ 4 \\ \hline \end{array}$$

a. 399
b. 391
c. 435
d. 398

III. Stimulus Attributes

1. The 1-digit number should be placed below the 3-digit number.

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

2. One of the distractors will be the answer if the student had subtracted. (b)

1. ADD WHOLE NUMBERS WITHOUT REGROUPING

I. General Description

Given an addition problem requiring no regrouping and consisting of a 3-digit number and a 2-digit number presented in a horizontal format, the student will select the correct answer from a list of alternatives.

II. Sample Item

40 + 831 + 428 =?

a. 1299
b. 1659
c. 1459
d. 1279

III. Stimulus Attributes

1. The two digit number should appear first in the presentation.

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

2. One distractors will be the result of locating the 2-digit number incorrectly in lining up the numbers in a vertical format. (b)

1. ADD WHOLE NUMBERS WITHOUT REGROUPING

I. General Description

Given an addition problem requiring no regrouping and consisting of two 3-digit numbers presented in a vertical format, the student will select the correct answer from a list of alternatives.

II. Sample Item

$$\begin{array}{r} 863 \\ \underline{+103} \end{array}$$

a. 966
b. 760
c. 960
d. 766

III. Stimulus Attributes

No other specifications.

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

2. One distractors will be the answer if the student had subtracted. (b)

2. ADD WHOLE NUMBERS WITH REGROUPING

I. General Description

Given an addition problem requiring regrouping and consisting of a 3-digit number and a 2-digit number presented in a vertical format, the student will select the correct answer from a list of alternatives.

II. Sample Item

$$\begin{array}{r} 783 \\ +\ 65 \\ \hline \end{array}$$

a. 848
b. 1148
c. 748
d. 728

III. Stimulus Attributes

1. The regrouping should occur in the tens to hundreds position.

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

2. One distractors will be the answer if the student mistakably carries the unit's digit. (b)

3. One distractors will be the answer if the student does not regroup. (c)

2. ADD WHOLE NUMBERS WITH REGROUPING

I. General Description

Given an addition problem requiring one regrouping and consisting of three 3-digit numbers presented in a vertical format, the student will select the correct answer from a list of alternatives.

II. Sample Item

$$\begin{array}{r} 450 \\ 645 \\ +866 \\ \hline \end{array}$$

a. 1961
b. 2461
c. 1851
d. 1861

III. Stimulus Attributes

No other specifications.

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

2. One distractors will be the answer if the student mistakably carries the unit's digit. (b)

3. One distractors will be the answer if the student does not regroup. (c)

2. ADD WHOLE NUMBERS WITH REGROUPING

I. General Description

Given an addition problem requiring two regrouping and consisting of three 3-digit numbers and a 2-digit number presented in a vertical format, the student will select the correct answer from a list of alternatives.

II. Sample Item

91 a. 1489
367 b. 2119
820 c. 1389
+211 d. 1419

III. Stimulus Attributes1

1. The 2-digit number should be the top number in the column.

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

2. One distractors will be the answer if the student mistakably carries the unit's digit. (b)

106

2. ADD WHOLE NUMBERS WITH REGROUPING

I. General Description

Given an addition problem requiring one regrouping and consisting of two 4-digit numbers presented in a vertical format, the student will select the correct answer from a list of alternatives.

II. Sample Item

$$\begin{array}{r} 3509 \\ +8812 \\ \hline \end{array}$$

a. 12321
b. 14121
c. 11311
d. 12311

III. Stimulus Attributes

1. The regroupings should occur in the units-to-tens position and the hundreds-to-thousands position.

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

2. One distractors will be the answer if the student mistakably carries the unit's digit. (b)

3. One distractors will be the answer if the student does not regroup. (c)

3. ADD DECIMALS

I. General Description

Given an addition problem involving two decimals of the form .x + .x presented in a vertical format, the student will select the correct answer from a list of alternatives.

II. Sample Item

$$\begin{array}{r} .6 \\ \underline{+.3} \end{array}$$

a. .9
b. 9
c. .09
d. .3

III. Stimulus Attributes

None

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

2. Distractors will include responses with the correct digits but with the decimal in the wrong place.

3. Decimal points will be aligned in the responses.

3. ADD DECIMALS

I. General Description

Given an addition problem involving two decimals of the form .x + x.x presented in a horizontal format, the student will select the correct answer from a list of alternatives.

II. Sample Item

.2 + 3.7 =?

a. 3.9
b. 39
c. 5.7
d. .39

III. Stimulus Attributes

None

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

2. Distractors will include responses with the correct digits but with the decimal in the wrong place.

3. Decimal points will be aligned in the responses.

4. One distractor will be the result of treating the .x number as if written in units. (c)

3. ADD DECIMALS

I. General Description

Given an addition problem involving two decimals of the form xx.x + xx.x presented in a vertical format, the student will select the correct answer from a list of alternatives.

II. Sample Item

$$\begin{array}{r} 21.8 \\ \underline{+16.9} \end{array}$$

a. 38.7
b. 3.87
c. 37.17
d. 37.7

III. Stimulus Attributes

None

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

2. Distractors will include responses with the correct digits but with the decimal in the wrong place.

3. Decimal points will be aligned in the responses.

3. ADD DECIMALS

I. General Description

Given an addition problem involving two decimals of the form $ .xx + $ .xx presented in a vertical format, the student will select the correct answer from a list of alternatives.

I. Sample Item

$ .66
+.71

a. $ 1.37
b. $ 137
c. $ .37
d. $ 13.70

III. Stimulus Attributes

1. Regrouping will occur from tenths to units position.

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.
2. Distractors will include responses with the correct digits but with the decimal in the wrong place.
3. Decimal points will be aligned in the responses.

3. ADD DECIMALS

I. General Description

Given an addition problem involving two decimals of the form \$x0.00 + \$xx.xx presented in a vertical format, the student will select the correct answer from a list of alternatives.

II. Sample Item

\$ 60.00	a.	\$123.91
63.19	b.	\$ 1.2391
	c.	\$ 12.39
	d.	\$123.81

III. Stimulus Attributes

1. The \$x0.00 number should be the top added.

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

2. Distractors will include responses with the correct digits but with the decimal in the wrong place.

3. Decimal points will be aligned in the response.

4. SUBTRACT WHOLE NUMBERS WITHOUT REGROUPING

I. General Description

Given a subtraction problem requiring no regrouping and consisting of two 2-digit numbers presented in a vertical format, the student will select the correct answer from a list of alternatives.

II. Sample Item

$$\begin{array}{r} 82 \\ \underline{-61} \end{array}$$

a. 21
b. 143
c. 141
d. 23

III. Stimulus Attributes

None

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

2. One of the distractors will be the answer the student would have obtained if added. (b)

4. SUBTRACT WHOLE NUMBERS WITHOUT REGROUPING

I. General Description

Given a subtraction problem requiring no regrouping and consisting of two 3-digit, (one including a zero) numbers presented in a vertical format, the student will select the correct answer from a list of alternatives.

II. Sample Item

$$\begin{array}{r} 864 \\ \underline{-202} \end{array}$$

a. 662
b. 1066
c. 666
d. 1062

II Stimulus Attributes

1. The zero should appear in the tens position of the second addend.

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

2. One of the distractors will be the answer the student would have obtained if added. (b)

4. SUBTRACT WHOLE NUMBERS WITHOUT REGROUPING

I. General Description

Given a subtraction problem requiring no regrouping and consisting of two 2-digit numbers presented in a vertical format, the student will select the correct answer from a list of alternatives.

II. Sample Item

769 - 48 =?

a. 721
b. 817
c. 801
d. 737

III. Stimulus Attributes

None

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

2. One of the distractors will be the answer the student would have obtained if added.

6. SUBTRACT DECIMALS

I. General Description

Given a subtraction problem involving two decimals of the form x.x - .x in a vertical format, the student will select the correct answer from a list of alternatives.

II. Sample Item

$$\begin{array}{r} 3.6 \\ \underline{-\ .3} \end{array}$$

a. 3.3
b. .3
c. .33
d. 33

III. Stimulus Attributes

1. There should be no regrouping required.

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct answer and three distractors.

6. SUBTRACT DECIMALS

I. General Description

Given a subtraction problem involving two decimals of the form x.x - .x in a horizontal format, the student will select the correct answer from a list of alternatives.

II. Sample Item

3.6 - .3 =?

a. 3.3
b. .6
c. .33
d. 33

III. Stimulus Attributes

1. There should be no regrouping required.

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct answer and three distractors.

6. SUBTRACT DECIMALS

I. General Description

Given a subtraction problem involving two decimals of the form \$x.xx -\$.xx in a vertical format, with regrouping the student will select the correct answer from a list of alternatives.

II. Sample Item

$$\begin{array}{r} \$\ 8.37 \\ -\quad .65 \\ \hline \end{array}$$

a. \$ 7.72
b. \$ 8.72
c. \$ 8.92
d. \$ 7.82

III. Stimulus Attributes

1. Regrouping will occur in the tenths-to-units position.

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct answer and three distractors.

2. One of the distractors will be the result of not regrouping. (b)

6. SUBTRACT DECIMALS

I. General Description

Given a subtraction problem involving two decimals of the form $xx.00 - $x.xx in a vertical format, the student will select the correct answer from a list of alternatives.

II. Sample Item

$$\begin{array}{r} \$\ 55.00 \\ -\ \underline{3.95} \end{array}$$

a. $ 51.05
b. $ 51.15
c. $ 52.05
d. $ 52.15

III. Stimulus Attributes

None

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct answer and three distractors.

6. SUBTRACT DECIMALS

I. General Description

Given a subtraction problem involving two decimals of the form $xx.95 - $x.xx in a horizontal format and requiring no regrouping, the student will select the correct answer from a list of alternatives.

II. Sample Item

$39.95 - $4.92 =? a. $35.03
b. $34.93
c. $35.83
d. $34.03

III. Stimulus Attributes

None

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct answer and three distractors.

6. SUBTRACT DECIMALS

I. General Description

Given a subtraction problem involving two decimals of the form .x - .x in a horizontal format, the student will select the correct answer from a list of alternatives.

II. Sample Item

.8 - .3 =?

a. .5
b. 5
c. .05
d. 5.5

III. Stimulus Attributes

None

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct answer and three distractors.

5. SUBTRACT WHOLE NUMBERS WITH REGROUPING

I. General Description

Given a subtraction problem requiring one regrouping and consisting of a 4-digit number and a 3-digit number presented in a vertical format, the student will select the correct answer from a list of alternatives.

II. Sample Item

$$\begin{array}{r} 6328 \\ -\ \underline{183} \end{array}$$

a. 145
b. 245
c. 211
d. 205

III. Stimulus Attributes

1. The regrouping should occur in the tens-to-hundreds position.

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

2. One of the distractors will be the result of not regrouping. (b)

5. SUBTRACT WHOLE NUMBERS WITH REGROUPING

I. General Description

Given a subtraction problem requiring two regrouping and consisting of two 3-digit numbers (one a multiple of 100) presented in a vertical format, the student will select the correct answer from a list of alternatives.

II. Sample Item

$$\begin{array}{r} 500 \\ -\underline{274} \end{array}$$

a. 226
b. 236
c. 274
d. 276

III. Stimulus Attributes

1. The multiple of one hundred will be the minuend.

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

5. SUBTRACT WHOLE NUMBERS WITH REGROUPING

I. General Description

Given a subtraction problem requiring one regrouping and consisting of two 4-digit numbers (one containing a zero) presented in a vertical format, the student will select the correct answer from a list of alternatives.

II. Sample Item

$$\begin{array}{r} 4068 \\ \underline{-2742} \end{array}$$

a. 1326
b. 2326
c. 2806
d. 1306

III. Stimulus Attributes

1. The zero should be located in the hundreds position of the minuends.

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

2. One of the distractors will be the result of not regrouping. (b)

5. SUBTRACT WHOLE NUMBERS WITH REGROUPING

I. General Description

Given a subtraction problem requiring two regrouping and consisting of two 4-digit numbers presented in a vertical format, the student will select the correct answer from a list of alternatives.

II. Sample Item

$$\begin{array}{r} 6043 \\ -\underline{5960} \end{array}$$

a. 83
b. 103
c. 183
d. 1083

III. Stimulus Attributes

1. The answer should be a 2-digit number.

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

5. SUBTRACT WHOLE NUMBERS WITH REGROUPING

I. General Description

Given a subtraction problem requiring two regrouping and consisting of two 3-digit numbers presented in a vertical format, the student will select the correct answer from a list of alternatives.

II. Sample Item

$$\begin{array}{r} 425 \\ \underline{-247} \end{array}$$

a. 178
b. 288
c. 188
d. 278

III. Stimulus Attributes

None

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

5. SUBTRACT WHOLE NUMBERS WITH REGROUPING

I. General Description

Given a subtraction problem requiring two regrouping and consisting of two 2-digit numbers presented in a vertical format, the student will select the correct answer from a list of alternatives.

II. Sample Item

62	a.	44
-18	b.	54
	c.	80
	d.	34

III. Stimulus Attributes

None

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

2. One of the distractors will be the result of not regrouping. (b)

7. MULTIPLY WHOLE NUMBERS WITHOUT CARRYING

I. General Description

Given a multiplication problem involving two 1-digit numbers the student will select the correct answer from a list of alternatives.

II. Sample Item

$$\begin{array}{r} 8 \\ \underline{\times 7} \end{array}$$

a. 56
b. 15
c. 42
d. 49

III. Stimulus Attributes

None

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

2. One distractor will be the answer the student would have obtained if the numbers were added. (b)

7. MULTIPLY WHOLE NUMBERS WITHOUT CARRYING

I. General Description

Given a multiplication problem involving a 1-digit and a 3-digit number but without carrying the student will select the correct answer from a list of alternatives.

II. Sample Item

$$\begin{array}{r} 602 \\ \times\ \ 3 \\ \hline \end{array}$$

a. 1806
b. 605
c. 186
d. 905

III. Stimulus Attributes

1. The multiplicand should have a zero in the tens position.

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

2. One distractor will be the answer the student would have obtained if the numbers were added. (b)

3. One distractor will be the answer the student would have obtained if the student had not used the correct place value. (c)

7. MULTIPLY WHOLE NUMBERS WITHOUT CARRYING

I. General Description

Given a multiplication problem involving a 2-digit and a 3-digit number but without carrying the student will select the correct answer from a list of alternatives.

II. Sample Item

276	a.	3036
x 11	b.	287
	c.	552
	d.	27876

III. Stimulus Attributes

1. The multiplier should be "11".

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

2. One distractor will be the answer the student would have obtained if the numbers were added. (b)

3. One distractor will be the answer the student would have obtained if the student had not used the correct place value. (c)

7. MULTIPLY WHOLE NUMBERS WITHOUT CARRYING

I. General Description

Given a multiplication problem involving two 2-digit numbers but without carrying the student will select the correct answer from a list of alternatives.

II. Sample Item

$$\begin{array}{r} 23 \\ \underline{\times 30} \end{array}$$

a. 690
b. 69
c. 53
d. 713

III. Stimulus Attributes

1. A zero should be in the units position of the multiplier.

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.
2. One distractor will be the answer the student would have obtained if the numbers were added. (c)
3. One distractor will be the answer the student would have obtained if the student had not used the correct place value. (b)

7. MULTIPLY WHOLE NUMBERS WITHOUT CARRYING

I. General Description

Given a multiplication problem involving two 1-digit numbers but without carrying the student will select the correct answer from a list of alternatives.

II. Sample Item

2 x 3 x 5 =?

a. 30
b. 10
c. 25
d. 45

III. Stimulus Attributes

1. Do not use zero as one of the factors.

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.
2. One distractor will be the answer the student would have obtained if the numbers were added. (b)

7. MULTIPLY WHOLE NUMBERS WITHOUT CARRYING

I. General Description

Given a multiplication problem involving a 2-digit and a multiple of 1000 in a horizontal format, the student will select the correct answer from a list of alternatives.

II. Sample Item

62 x 4000 =?

a. 248000
b. 28000
c. 242000
d. 32000

III. Stimulus Attributes

None

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

7. MULTIPLY WHOLE NUMBERS WITHOUT CARRYING

I. General Description

Given a multiplication problem involving two 2-digit numbers but without carrying the student will select the correct answer from a list of alternatives.

II. Sample Item

$$\begin{array}{r} 50 \\ \times 40 \\ \hline \end{array}$$

a. 2000
b. 200
c. 90
d. 2050

III. Stimulus Attributes

1. Zeroes should be in the units position of both numbers.

IV. Response Attributes

1. Students will be asked to mark the letter of one of four given response alternatives consisting of the correct response and three distractors.

2. One distractor will be the answer the student would have obtained if the numbers were added. (c)

3. One distractor will be the answer the student would have obtained if the student had not used the correct place value. (b)

8. MULTIPLY WHOLE NUMBERS WITH CARRYING

I. General Description

Given a multiplication problem involving a 2-digit and a 3-digit number and with carrying, the student will select the correct answer from a list of alternatives.

II. Sample Item

718	a.	62466
x 87	b.	805
	c.	10670
	d.	5676

III. Stimulus Attributes

None

IV. Response Attributes

1. Students will be asked to mark the better of one of four given response alternatives consisting of the correct response and three distractors.

8. MULTIPLY WHOLE NUMBERS WITH CARRYING

I. General Description

Given a multiplication problem consisting of three 1-digit numbers in a horizontal format, the student will select the correct answer from a list of alternatives.

II. Sample Item

6 x 7 x 5 =?

a. 210
b. 225
c. 425
d. 47

III. Stimulus Attributes

None

IV. Response Attributes

1. Students will be asked to mark the better of one of four given response alternatives consisting of the correct response and three distractors.

8. MULTIPLY WHOLE NUMBERS WITH CARRYING

I. General Description

Given a multiplication problem involving a 2-digit and a 1-digit and with carrying, the student will select the correct answer from a list of alternatives.

II. Sample Item

$$\begin{array}{r} 38 \\ \times\ \underline{4} \end{array}$$

a. 152
b. 42
c. 132
d. 102

III. Stimulus Attributes

None

IV. Response Attributes

1. Students wil be asked to mark the better one of four given response alternatives consisting of the correct response and three distractors.

8. MULTIPLY WHOLE NUMBERS WITH CARRYING

I. General Description

Given a multiplication problem involving two 2-digit numbers and with carrying, the student will select the correct answer from a list of alternatives.

II. Sample Item

$$\begin{array}{r} 57 \\ \underline{\times 42} \end{array}$$

a. 2394
b. 99
c. 342
d. 2014

III. Stimulus Attributes

None

IV. Response Attributes

1. Students will be asked to mark the better of one of four given response alternatives consisting of the correct response and three distractors.

8. MULTIPLY WHOLE NUMBERS WITH CARRYING

I. General Description

Given a multiplication problem involving a 1-digit and a 3-digit number and with carrying, the student will select the correct answer from a list of alternatives.

II. Sample Item

$$\begin{array}{r} 846 \\ \times\ \underline{\quad 7} \end{array}$$

a. 5922
b. 853
c. 5682
d. 842

III. Stimulus Attributes

None

IV. Response Attributes

1. Students will be asked to mark the better of one of four given response alternatives consisting of the correct response and three distractors.

8. MULTIPLY WHOLE NUMBERS WITH CARRYING

I. General Description

Given a multiplication problem involving a 1-digit and a 3-digit number and with carrying, the student will select the correct answer from a list of alternatives.

II. Sample Item

804	a.	5628
x 7	b.	811
	c.	5608
	d.	588

III. Stimulus Attributes

1. The multiplicand should contain a zero in the tens position.

IV. Response Attributes

1. Students will be asked to mark the better of one of four given response alternatives consisting of the correct response and three distractors.

9. DIVIDE WHOLE NUMBERS WITHOUT REMAINDERS

I. General Description

Given a division problem involving basic division facts through 81 9 in a "   " format that will not result in a remainder, the student will select the correct answer from a list of alternatives.

II. Sample Item

72 ÷ 8 =?

a. 9
b. 80
c. 64
d. 576

III. Stimulus Attributes

None

IV. Response Attributes

1. Students will be asked to mark the better of one of four given response alternatives consisting of the correct response and three distractors.

9. DIVIDE WHOLE NUMBERS WITHOUT REMAINDERS

I. General Description

Given a division problem involving 2-digit number divided by a 1-digit number with regrouping in a "   " format that will not result in a remainder, the student will select the correct answer from a list of alternatives.

II. Sample Item

4 68

a. 17
b. 64
c. 24
d. 27

III. Stimulus Attributes

None

IV. Response Attributes

1. Students will be asked to mark the better of one of four given response alternatives consisting of the correct response and three distractors.

9. DIVIDE WHOLE NUMBERS WITHOUT REMAINDERS

I. General Description

Given a division problem involving 3-digit number divided by a 2-digit number in a "   " format that will not result in a remainder, the student will select the correct answer from a list of alternatives.

II. Sample Item

24 864

a. 36
b. 81
c. 63
d. 18

III. Stimulus Attributes

None

IV. Response Attributes

1. Students will be asked to mark the better of one of four given response alternatives consisting of the correct response and three distractors.

9. DIVIDE WHOLE NUMBERS WITHOUT REMAINDERS

I. General Description

Given a division problem involving 2-digit number divided by a 2-digit number in a "   " format that will not result in a remainder, the student will select the correct answer from a list of alternatives.

II. Sample Item

90 18 =?

a. 5
b. 6
c. 12
d. 72

III. Stimulus Attributes

None

IV. Response Attributes

1. Students will be asked to mark the better of one of four given response alternatives consisting of the correct response and three distractors.

9. DIVIDE WHOLE NUMBERS WITHOUT REMAINDERS

I. General Description

Given a division problem involving 3-digit number divided by a 2-digit number in a "   " format that will not result in a remainder, the student will select the correct answer from a list of alternatives.

II. Sample Item

20 180

a. 9
b. 90
c. 160
d. 15

III. Stimulus Attributes

1. The divisor will be a multiple of ten.

IV. Response Attributes

1. Students will be asked to mark the better of one of four given response alternatives consisting of the correct response and three distractors.

MATH TEST

DIRECTIONS

Do each of the problems on the following pages and select the correct answer for each. Then mark the space on your answer sheet with the same as letter as the answer you select. You may mark on this test, but don't put any stray marks on your answer sheet.

SAMPLE ITEM

$$\begin{array}{r} 18 \\ +\underline{21} \end{array}$$

A. 12
B. 39
C. 42
D. 20

On your answer sheet you should fill in the space corresponding to "B" because 39 is the correct answer. Your answer sheet would look like this:

⊂A⊃ ● ⊂C⊃ ⊂D⊃

Use a #2 pencil to mark you answers. If you make any changes, be sure to erase completely.

IS

1

395
+ 4

A. 435
B. 399
C. 391
D. 398

2

863
+103

A. 960
B. 966
C. 760
D. 766

3

40 + 831 + 428 =

A. 1,459
B. 1,279
C. 1,299
D. 1,659

4

793
+ 45

A. 738
B. 748
C. 848
D. 838

5

6,549
+4,815

A. 11,364
B. 10,354
C. 13,394
D. 11,354

6

450
645
+866

A. 1,851
B. 1,961
C. 1,861
D. 2,461

7

91
367
820
+211

A. 2,119
B. 1,419
C. 1,489
D. 1,389

8

.6
+ .3

A. .9
B. 9
C. .09
D. .3

9

.2 + 3.7 =

A. .39
B. 5.7
C. 39.
D. 3.9

10

21.8
+16.9

A. 3.87
B. 37.17
C. 38.7
D. 37.7

11

$ 60.00
+63.19

A. $ 12.319
B. $123.09
C. $123.19
D. $120.19

12

$ .66
+.71

A. $13.70
B. $ .37
C. $1.37
D. $137.00

13

82
-61

A. 23
B. 141
C. 143
D. 21

14

864
-202

A. 662
B. 1,066
C. 666
D. 1,062

15

52
-19

A. 71
B. 33
C. 43
D. 34

16

769 - 48 =

A. 817
B. 801
C. 737
D. 721

17

6,328
- 183

A. 6,245
B. 6,211
C. 6,205
D. 6,145

18

4,068
-2,742

A. 2,326
B. 1,326
C. 2,806
D. 1,306

19

500 - 274 =

A. 236
B. 326
C. 226
D. 374

20

6,043
-5,960

A. 83
B. 1,083
C. 183
D. 103

21

$$\begin{array}{r} 425 \\ \underline{-247} \end{array}$$

A. 288
B. 178
C. 188
D. 278

22

$$\begin{array}{r} 3.6 \\ \underline{-\ .3} \end{array}$$

A. .3
B. 33
C. .33
D. 3.3

23

6.7 - .5 =

A. .62
B. .2
C. 6.2
D. 62

24

.8 - .3 =

A. .05
B. .5
C. 5
D. 5.5

25

$$\begin{array}{r} \$\ 8.37 \\ \underline{-\ .65} \end{array}$$

A. $7.72
B. $7.82
C. $8.92
D. $8.72

26

$39.95 - $4.92 =

A. $35.83
B. $35.03
C. $34.03
D. $34.93

27

$$\begin{array}{r} \$\ 55.00 \\ \underline{-\ 3.95} \end{array}$$

A. $51.05
B. $51.15
C. $52.05
D. $52.15

28

$$\begin{array}{r} 8 \\ \underline{\times 7} \end{array}$$

A. 42
B. 56
C. 49
D. 15

29

$$\begin{array}{r} 38 \\ \underline{\times\ 4} \end{array}$$

A. 102
B. 152
C. 42
D. 132

30

$$\begin{array}{r} 602 \\ \underline{\times\ \ 3} \end{array}$$

A. 186
B. 1,806
C. 605
D. 1,836

31

$$\begin{array}{r} 23 \\ \times\underline{30} \end{array}$$

A. 690
B. 69
C. 53
D. 713

32

$$\begin{array}{r} 409 \\ \times\underline{\ \ 7} \end{array}$$

A. 343
B. 2,803
C. 2,863
D. 416

33

50 x 40 =

A. 90
B. 200
C. 2,000
D. 2,090

34

$$\begin{array}{r} 276 \\ \times\underline{\ 11} \end{array}$$

A. 27,876
B. 522
C. 3,036
D. 287

35

2 x 3 x 5 =

A. 10
B. 25
C. 30
D. 45

36

6 x 7 x 5 =

A. 425
B. 225
C. 210
D. 47

37

$$\begin{array}{r} 57 \\ \times\underline{42} \end{array}$$

A. 2,394
B. 342
C. 2,014
D. 99

38

$$\begin{array}{r} 846 \\ \times\underline{\ \ 9} \end{array}$$

A. 7,614
B. 855
C. 1,723
D. 7,254

39

62 x 4,000 =

A. 248,000
B. 242,000
C. 32,000
D. 28,000

40

$$\begin{array}{r} 518 \\ \times\underline{\ 37} \end{array}$$

A. 5,180
B. 15,156
C. 555
D. 19,166

41

72 ÷ 8 =

A. 64
B. 80
C. 7
D. 9

42

$4\overline{)68}$

A. 12
B. 27
C. 24
D. 17

43

$24\overline{)864}$

A. 18
B. 36
C. 41
D. 32

44

$20\overline{)180}$

A. 9
B. 90
C. 81
D. 15

45

90 ÷ 18 =

A. 5
B. 12
C. 6
D. 20

Reading Test

DIRECTIONS

Read each story and then answer the items following it. Choose the best answer for each item. Then mark the space on your answer sheet with the same letter as the answer you select.

SAMPLE STORY

Susan was very sleepy. The family had been riding in the car for over five hours. She would be very glad when they got home. Then she could sleep in her own bed.

SAMPLE ITEM

Where was Susan?

A. In bed
B. In a bus
C. In a car
D. In an office

You should have filled in the space "C" on your answer sheet because Susan was in a car. Your answer sheet would look like this:

⊂A⊃ ⊂B⊃ ● ⊂D⊃

Use a #2 pencil to mark your answers.

If you make any changes, be sure to erase completely.

IB

"Going to the mountains is going home". This was written by John Muir. Muir was a quite man who loved the mountains. Much of his life was spent trying to preserve the wilderness.

Muir was born in Scotland. In 1868 he moved to California where he lived the rest of his life. Muir loved to spend weeks in the California mountains. He would carry only his notebook, fried bread and tea.

Muir felt that people must have wilderness areas. He convinced many people there was a need. His writing about the wilderness resulted in the establishment of three national parks.

Muir died in 1914, but he is still remembered today. The Sierra Club, a group he founded in 1892, is still devoted to saving the wilderness. Many people continue to read the books he wrote. After his death, a trail was developed and named for him, the John Muir Trail.

---

1. John Muir was born in

   A. California.
   B. the mountains of Ohio.
   C. Scotland.
   D. the Sierra Nevada Mountains.

2. What caused the establishment of three national parks?

   A. Muir started the Sierra Club.
   B. John Muir Trail was created.
   C. Muir moved to California.
   D. Muir wrote about the wilderness.

3. How do we know John Muir loved the wilderness?

   A. He moved to California.
   B. He was a quiet man.
   C. People still read his books.
   D. He spent weeks in the mountains.

4. Which happened last in the story?

   A. The Sierra Club was formed.
   B. Muir moved to California.
   C. Muir died.
   D. The John Muir Trail was developed.

5. When was the Sierra Club formed?

   A. 1868
   B. 1892
   C. 1914
   D. 1920

6. The best title for this story is:

   A. The Life of John Muir
   B. Three National Parks
   C. Born in Scotland
   D. Fried Bread and Tea

Read each sentence. If it tells a fact, choose A. If it tells something that is just a matter of how people think about it, an opinion, choose B.

---

7. There are more women over 70 than men.

   A. fact
   B. opinion

8. In Summer we usually have warmer weather than in winter.

   A. fact
   B. opinion

9. There are 26 letters in the alphabet.

   A. fact
   B. opinion

10. According to the legend, Johnny Appleseed wore a cooking utensil for a hat.

   A. fact
   B. opinion

Father's Day was approaching. Emily and Tom wanted to surprise their Dad. They did not have any money for a present, though. "Let's make Dad some brownies," said Emily. "He really likes chocolate, and Mom rarely makes dessert."

"Brownies taste good," said Tom. "I like them a lot."

Tom found a cookbook, and they looked up a recipe. Emily gathered the ingredients. Tom got out bowls, pans, and a mixer. Soon the brownies were mixed and in the oven. Tom set the timer. After 10 minutes, Emily said, "Oh, no, I forgot baking soda. That's what makes the brownies rise!"

---

11. Why did Tom and Ellen think Dad would like brownies?

A. Tom liked brownies.
B. Father's Day was approaching.
C. Dad likes chocolate.
D. Tom found a cookbook.

12. Who found the cookbook?

A. Mom
B. Dad
C. Tom
D. Emily

13. What happened before Emily found the ingredients?

A. Tom found a cookbook.
B. They mixed the brownies.
C. Tom set the timer.
D. Tom got out bowls.

14. Tom and Ellen are

A. friends.
B. parents.
C. cousins.
D. brother and sister.

15. The best title for this passage is:

A. A Father's Day Present
B. Tom Likes Brownies
C. How to Make Brownies
D. Mom Makes A Dessert

16. Which statement is an opinion?

A. Father's Day was approaching.
B. Brownies taste good.
C. Tom found a cookbook.
D. Oh, no, I forgot baking soda.

17. The next time they bake brownies, what should Tom and Ellen do differently?

A. Make more brownies.
B. Double check the ingredients.
C. Save money for a present.
D. Have Mom make more desserts.

18. When Tom and Emily take the brownies out of the oven, they will be

A. high and fluffy.
B. flat.
C. gooey.
D. sugar free.

We usually think of sugar coming from sugar cane, but it is also made from sugar beets. The sugar beet is an important crop. It is grown in some of our Western states. When the beets are fully grown, the farmers dig them out of the ground. Then they load the beets on trucks and take them to a sugar factory.

Do you know how we get sugar from the sugar beet? At the factory, the beets are washed, sliced, and then placed in hot water. The sugar in the beet chips dissolves in the water. Clear sugar water is drained off and placed in ovens. There it cooks into a thick syrup. Then the syrup goes into large pans. The syrup is boiled until sugar begins to form.

Then the syrup is placed in the bowl on a machine that spins very fast. The liquid is removed, and sugar is left on the sides of the bowl. The sugar is then placed on driers. The driers remove any remaining liquid.

---

19. What happens first to the beets at the factory?

A. The beets are placed in hot water.
B. The beets are dug out of the ground.
C. The sugar water from the beets is placed in ovens.
D. The beets are washed.

20. What is the last thing that is done in the factory to make sugar?

A. The syrup is spun in a machine.
B. The beets are loaded on a truck.
C. The remaining liquid is removed.
D. The sugar water is cooked.

21. Where are sugar beets grown?

A. South America
B. in our Western states
C. in our Eastern states
D. Hawaii

22. This story is about

A. where the sugar beet is grown.
B. the machinery used to make beets into sugar.
C. how sugar is spun from beet syrup.
D. how sugar beets become sugar.

23. The author wants the reader to understand

A. the importance of water in the process of making sugar.
B. farmers have to work hard to make sugar from beets.
C. we wouldn't have any sugar if it weren't for sugar beets.
D. trains are important transportation in the sugar beet industry.

There was a king who liked to sit outside and look at the North Star. One night when the star was very bright, the king said, "I must go there. I will think of a way."

The next day the king sent for the best builder in his kingdom. "I want you to build a tower to the North Star," said the king.

The builder asked for boxes of all sizes. His helpers put one box on top of another until the pile was very high.

Then the king began to climb the tower. He climbed higher and higher until he came to the top. But the top of the tower did not reach the star. The king needed one more box. So the builders took a box from the bottom of the pile.

---

24. Which of these things happened first?

    A. The king climbed the tower.
    B. The king called his builder.
    C. The king looked at the North Star.
    D. The king needed another box.

25. What probably happened at the end of the story?

    A. The king got to the North Star.
    B. The builder finished the tower.
    C. The boxes were painted.
    D. The tower and the king fell down.

26. In the third paragraph, His is underlined.
    Who does His refer to?

    A. the builder
    B. the king
    C. both the builder and the king
    D. the builder's helpers

27. In the fourth paragraph, he is underlined.
    Who does he refer to?

    A. the builder
    B. the king
    C. the builder's helpers
    D. the box maker

28. This is a story about a king who

    A. wanted to touch the sky.
    B. had a dream about reaching the North Star.
    C. wanted to visit the North Star.
    D. reached the North Star.

Johnny Appleseed trudged barefooted through the frontier with a burlap bag, holding a treasure of more value than gold, over his shoulder. He carried apple seeds, enough to provide all the territory with trees. The apple was not native to this land, and the seeds had to be imported and planted carefully. That was Johnny's mission--to plant thriving apple orchards throughout the Ohio Valley.

Johnny started with a nursery of his own in Pittsburgh. Then he planted seeds wherever they would grow. He worked deeper and deeper into the frontier country to land where no white man had ever been. So when people moved west and claimed their farms, they found thriving apple orchards waiting for them.

---

29. Choose the best ending for this sentence:

    The first appleseeds had to be imported because

    A. the apple trees kept dying.
    B. nobody lived deep in the frontier country.
    C. the apple did not originally come from America.
    D. apples used to be very small.

30. Choose the best conclusion for this passage:

    A. Johnny Appleseed wore a coonskin cap in cold weather.
    B. No apple seeds were planted in the unsettled country.
    C. Johnny Appleseed performed a great service for the settlers in the Ohio Valley.
    D. White men lived in all the country to which Johnny Appleseed traveled.

31. Which of the following statements is true?

    A. Johnny Appleseed was an Indian.
    B. Pittsburgh is east of the Ohio Valley.
    C. People who moved into the Ohio Valley were farmers from Pittsburgh.
    D. Farmers hired Johnny Appleseed to plant apple orchards for them.

32. When did Johnny Appleseed begin planting apple seeds?

    A. After he traveled to Europe.
    B. As soon as the apple seeds were sent.
    C. Once he had apple seeds from his nursery.
    D. After an invitation from the frontier pioneers.

The common use of airplanes for travel is fairly new. Airplanes were developed in the early 1900's. They were used as fighters in the two world wars. Until 1950, however, very few people used them for travel. Even then, the planes were small and flying was uncomfortable. Now that we have large, comfortable jets, flying is common. However, it would be fun to fly in an old airplane to see how people used to travel.

---

33. This is mainly a story about

A. jet airplanes.
B. airplanes being used for transportation.
C. uses of airplanes in world wars.
D. the design of comfortable airplanes.

34. In the 1950's, planes were small and uncomfortable. What effect did this have?

A. Planes were used as fighters.
B. Jets were developed.
C. Few people traveled by plane.
D. Many people traveled by plane, but they were uncomfortable.

35. The article states that we now have large, comfortable jets. What result does this have?

A. Airplanes were developed.
B. The author would like to fly in an old airplane.
C. People buy airplanes.
D. Flying is common.

36. Which statement is an opinion?

A. Planes were used as fighters.
B. Now we have large jets.
C. Until 1950 very few people used planes for travel.
D. It would be fun to fly in an old airplane.

37. In the fourth sentence, them is underlined. To what does the word refer?

A. fighter airplanes
B. pilots
C. airplanes
D. jet airplanes

As the cat climbed over
the button box
He went one foot at a time

The right forepaw
    came
        down
            safely

But a hind foot caught
In the button box

Buttons
Scattered in all directions.

---

38. Which of the following sentences tells what probably happened after the buttons were scattered?

A. The cat ran away from the box and buttons.
B. The cat picked up the buttons.
C. The cat fell asleep.
D. The cat fell and hurt himself.

When the first people settled in the Mohawk River Valley the river was clean and pure. They started towns at a number of places and used the river as a transportation route for people and products.

In the late 1800's and early 1900's, factories were started in these towns. People and factories often have a lot of waste materials to get rid of. Finding a place to put them can be a great problem. Towns and factories along the river answered this problem by dumping wastes into the river. As time went on, more and more wastes were dumped into the river. Some of these wastes were chemicals that were harmful to the animal and plant life existing in the river. The river became dirtier and dirtier until the water was no longer fit to drink; few fish could live in it; and it was not safe for people to swim in. Now that people have realized this, they are attempting to stop dirtying the river so that it is once again fit to use.

---

39. Which of the following sentences best states the main idea of this passage?

    A. When the first settlers made their homes along the Mohawk River, it was clean and pure.
    B. Our rivers are being kept clean and pure.
    C. No one realizes what made the Mohawk River dirty.
    D. The Mohawk River has become unfit for plant and animal life and harmful for people to use.

40. According to the passage, the present condition of the Mohawk River is a result of

    A. the activities of the Indians who used the river for transportation before the white man came.
    B. the fish dying in the river and making it dirty.
    C. too many people using the river for transportation.
    D. man using the river as a dumping area for wastes.

41. The Mohawk River was a clean river in the 1800's because

    A. it was a young river.
    B. there were strong anti-pollution laws.
    C. there were no people or towns around to pollute the river.
    D. chemicals hadn't been invented yet.

42. What does the story indicate will be necessary to stop pollution of the Mohawk River?

    A. Anti-pollution laws and efforts must be made.
    B. Factories will have to move to other areas, away from the river.
    C. The Valley will have to be abandoned.
    D. There will have to be a lot of picketing.

43. The author believes

    A. rivers would remain clean if factories stayed away.
    B. rivers can be kept clean and safe if wastes are not dumped into them.
    C. fish cannot live in rivers where towns are nearby.
    D. modern factories pollute the river and the air.

The old man, a Greek, fought the Roman army to a standstill for nearly 3 years—and almost won. The old man was Archimedes of Syracuse, the greatest scientist of the ancient world.

The Roman army knew his reputation well, and he lived up to it fully. Legend says that when curved mirrors were set up on the walls of Syracuse, a Greek city in Sicily, the besieging Roman ships caught fire. It wasn't sorcery; it was Archimedes. When huge claws were extended outward on beams, ships were caught, raised and overturned. It wasn't magic; it was Archimedes.

---

44. According to the passage, which of the following would most cause the ships to burn?

A. Strong winds caused the boats to capsize.
B. Huge claws caused the ships to burn.
C. The Roman ships caught fire because of reflecting mirrors.
D. Archimedes hypnotized the sailors.

45. If Archimedes were alive today, he would probably be a

A. movie producer.
B. writer.
C. aerospace scientist.
D. gardener.

162

MATH TEST

DIRECTIONS

Do each of the problems on the following pages and select the correct answer for each. Then mark the space on your answer sheet with the same as letter as the answer you select. You may mark on this test, but don't put any stray marks on your answer sheet.

SAMPLE ITEM

$$\begin{array}{r}18\\+\underline{21}\end{array}$$

A. 12
B. 39
C. 42
D. 20

On your answer sheet you should fill in the space corresponding to "B" because 39 is the correct answer. Your answer sheet would look like this:

⊏A⊐ ■B■ ⊏C⊐ ⊏D⊐

Use a #2 pencil to mark you answers. If you make any changes, be sure to erase completely.

IB

1

12
+13

A. 25
B. 26
C. 35
D. 36

2

40 + 831 + 428 =

A. 1,459
B. 1,279
C. 1,299
D. 1,659

3

7,418
361
+2,100

A. 5,879
B. 9,879
C. 9,889
D. 9,979

4

37
+85

A. 102
B. 112
C. 122
D. 123

5

834
+166

A. 900
B. 990
C. 1,000
D. 1,100

6

91 + 367 +820+211 =

A. 2,119
B. 1,419
C. 1,489
D. 1,389

7

4,218
7,564
+3,211

A. 13,893
B. 14,883
C. 14,893
D. 14,993

8

94
-21

A. 63
B. 73
C. 75
D. 115

9

864
-202

A. 662
B. 1,066
C. 666
D. 1,062

10

769 - 48 =

A. 817
B. 801
C. 737
D. 721

11

$$\begin{array}{r} 41 \\ \underline{-25} \end{array}$$

A. 15
B. 16
C. 25
D. 26

12

$$\begin{array}{r} 525 \\ \underline{-416} \end{array}$$

A. 109
B. 111
C. 131
D. 141

13

6,328 - 183 =

A. 6,245
B. 6,211
C. 6,205
D. 6,145

14

$$\begin{array}{r} 380 \\ \underline{-295} \end{array}$$

A. 85
B. 95
C. 105
D. 115

15

500 - 274 =

A. 235
B. 326
C. 226
D. 374

16

$$\begin{array}{r} 9{,}841 \\ \underline{-2{,}888} \end{array}$$

A. 6,953
B. 6,963
C. 7,053
D. 7,953

17

$$\begin{array}{r} 21.8 \\ \underline{+16.9} \end{array}$$

A. 3.87
B. 37.17
C. 38.7
D. 37.7

18

$$\begin{array}{r} \$\ 60.00 \\ \underline{+63.19} \end{array}$$

A. $ 12.319
B. $123.09
C. $123.19
D. $120.19

19

.0554 + .7005 =

A. .7559
B. 1.7559
C. 7.559
D. 7559

20

$$\begin{array}{r} 5.2283 \\ \underline{+81.261} \end{array}$$

A. 6.0409
B. 6.0509
C. 86.4893
D. 86.5893

21

835.06 + 178.1 =

A. 1,013.07
B. 1,013.16
C. 8,428.7
D. 8,528.7

22

3.6
- .3

A. .3
B. 33
C. .33
D. 3.3

23

$ 8.37
- .65

A. $7.72
B. $7.82
C. $8.92
D. $8.72

24

6.653 - 3.43 =

A. 3.123
B. 3.223
C. 3.31
D. 6.310

25

34.48
-33.923

A. 0.125
B. 0.563
C. 0.443
D. 0.557

26

1.3568 - 1.254 =

A. .0028
B. .1028
C. 1.2314
D. 12.314

27

176.3 - 2.15 =

A. 18.48
B. 17.415
C. 154.8
D. 174.15

28

8 x 7 =

A. 1
B. 15
C. 56
D. 57

29

7 x 0 =

A. 0
B. 2
C. 7
D. 14

30

32
x 4

A. 76
B. 126
C. 128
D. 138

31

$$\begin{array}{r}602\\ \times\ \underline{\quad 3}\end{array}$$

A. 186
B 1,806
C. 605
D. 1,836

32

50 x 40 =

A. 90
B. 200
C. 2,000
D. 2,090

33

$$\begin{array}{r}23\\ \times\underline{30}\end{array}$$

A. 690
B. 69
C. 53
D. 713

34

$$\begin{array}{r}276\\ \times\ \underline{11}\end{array}$$

A. 27,876
B. 522
C. 3,036
D. 287

35

$$\begin{array}{r}65\\ \times\ \underline{4}\end{array}$$

A. 240
B. 249
C. 260
D. 280

36

$$\begin{array}{r}38\\ \times\ \underline{4}\end{array}$$

A. 102
B. 152
C. 143
D. 132

37

$$\begin{array}{r}409\\ \times\ \underline{\quad 7}\end{array}$$

A. 343
B. 2,803
C 2,863
D. 416

38

$$\begin{array}{r}57\\ \times\underline{42}\end{array}$$

A. 2,394
B. 342
C. 2,014
D. 99

39

$$\begin{array}{r}15\\ \times\underline{78}\end{array}$$

A. 630
B. 930
C. 1,170
D. 1,270

40

$$\begin{array}{r}518\\ \times\ \underline{37}\end{array}$$

A. 5,180
B. 15,156
C. 555
D. 19,166

41

72 ÷ 8 =

A. 64
B. 80
C. 7
D. 9

42

$4\overline{)68}$

A. 12
B. 27
C. 24
D. 17

43

783 ÷ 9 =

A. 84
B. 85
C. 86
D. 87

44

90 ÷ 18 =

A. 5
B. 12
C. 6
D. 20

45

$35\overline{)3{,}535}$

A. 11
B. 101
C. 110
D. 111

October 9, 1981

Marilyn Averill
Evaluation Specialist
Boulder Valley School District
P.O. Box 9011
Boulder, Colorado 80301

Dear Marilyn:

Here, finally, is the information I told you I would send regarding the NIE study I am working on for the Northwest Regional Educational Laboratory. If you need or would like additional information, don't hesitate to call. You will see in the packet of information that I say I will send 100 questionnaires to you (25 copies of 4 different questionnaires). While it would be wonderful to actually get that many responses from Boulder, I know that may be impossible because you may not even have that many grade 3-6 teachers and reading and math specialists. Obviously, the more responses I can get, the better. Your help will be very much appreciated, no matter how many respondents from Boulder I eventually end up with.

I will be mailing the questionnaires to you late during the week of October 19. If anything will prevent you or your teachers from participating in this research activity, would you let me know by that Monday (October 19)? If I don't hear from you or someone else in your office, I'll go on and mail the questionnaires.

Thanks again for your willingness to help me out. If I can return the favor, be sure to let me know. Hope your year is off to a good start and that things go well for you this year.

Best wishes,

Evelyn J. Brzezinski
Senior Associate

Attachment
EJB/hs

AN EMPIRICAL INVESTIGATION OF TWO TEST DEVELOPMENT STRATEGIES

Summary of the Study

The following information is provided in response to questions about the study posed by one school district in which teachers are being asked to respond to questionnaires. If additional information is desired, please do not hesitate to contact Evelyn Brzezinski, the study's principal investigator, at (503) 223-3396.

Benefits of the Study to the School District and Education in General

With the increasingly tight funding situation facing education, administrators and testing specialists are looking for ways to get the most pay-off from their test development dollars. The use of item banks is viewed by many as a promising direction because it appears to obviate the need for the very time consuming and expensive steps of developing item specifications (which give detailed requirements for prototype items) and then, from them, the items themselves. But Popham, Hambleton and others tell us that those steps are critical to the development of high quality objective referenced tests (ORTs). According to those experts, item specifications are necessary to ensure high content validity of the tests, while item development activities yield a feeling of ownership (which leads, presumably, to increased use) on the part of the persons who develop the tests. But as long as people feel compelled to carry out both steps, the maximum utility of item banks cannot be realized.

What is needed is research to determine whether the item specification and item development steps are, in fact, necessary in order to produce ORTs with high content validity and teacher acceptance. This study does just that. It contrasts two different methods for test development (item specification based and item bank based) and looks at the cost and quality of the products those methods produce. If there are differences in the quality of the tests, the magnitude of the differences will be compared with the differences in cost to see which procedure to recommend. If it turns out that the tests have comparable quality but that the item bank method is less expensive, test developers around the country can begin to take more advantage of item collections (as Boulder Valley, Central Valley, Grand Rapids, Portland and other school districts have done.) And they can take advantage of existing item collections knowing that the resulting tests will be as good as if they had been developed from scratch.

Research Questions to be Explored

1. Is there a difference in the psychometric quality of tests developed by two different methods (item specifications and item bank)?

2. Is there a difference in the content validity of tests developed by the two methods?

3. Is there a difference in the quality of tests (as perceived by teachers) developed by the two methods?

4. Is there a difference in the cost of tests developed by the two methods?

5. Based on answers to the above questions, is one test development method to be

recommended over the other?

Brief Summary of Related Literature

Much literature exists on how to develop good objective referenced tests, and there is considerable consistency in the recommended procedures. But that literature presumes that tests are being developed from scratch: content is being specified and items are being written to measure the desired content. Test development using item banks, on the other hand, is an area with very little research. An ERIC search crossing ITEM BANKS with COMPUTERS and ITEM BANKS with TEST CONSTRUCTION resulted in 126 citations. Professional writings on the subject of item banks seem to be divided into two major categories: either (1) how to organize a collection of items for item banking purposes or (2) how a particular item bank should be or has been used in practice. A review of the 126 documents failed to produce any study which looked at the quality of tests produced from an item bank compared to the quality of tests developed from scratch. Thus, apparently people are using item banks to develop tests but do not know (or at least do not report) how the quality of those tests compares with other tests developed in the way recommended by Popham, Hambleton and others. No one seems to be investigating how test development procedures should be adapted from current standard practice in order to take advantage of the fact that, with access to banks, items are already available and do not have to be written. Clearly, this is a tremendous gap in the body of knowledge related to test development. If item banks are to achieve their full potential, it is a gap which must be closed by research studies such as the one described here.

Study Design

Content panels composed of teachers and curriculum specialists in the Portland, Oregon area will be formed and charged with the task of developing tests based on a specified content listing provided by the principal investigator. Two panels (one in reading, one in math) will work with a testing specialist from the Northwest Regional Educational Laboratory (NWREL) to develop item specifications and sample items, and review resulting 45-item end-of-fourth-grade reading and math tests. Two other panels (again, one in reading and one in math) with go directly from the same content listings to a subsample of the 20,000-item NWREL basic skills item bank. They will review items in the categories specified by the content listing and select items to go on the 45-item end-of-fourth-grade reading and math tests. One third of the items those panelists review will be the items which appear on the first content panels' tests.

Independent variables in the study are the time and cost required to develop each test. Dependent variables include technical quality of the tests, content validity of the tests, and perceived quality of the tests. The technical quality of the tests will be determined by administering each test to approximately 75-100 fifth grade students in Central Valley (WA) School District in October 1981. Each test (item specification-produced and item bank-produced) will be given to a random half of enough classrooms to complete the sample size. Item analyses will be conducted and t-tests on the differences between test means, standard errors and reliabilities will be computed.

The content validity variable will be measured using volunteer upper elementary teachers and curriculum experts in several school districts in Colorado, Michigan and Oregon. Each reviewer will receive a questionnaire (see Attachment A) which lists all 90 items from the two 45-item reading tests (or math tests, depending on the reviewer's area of expertise) and the content listing from which the tests were developed. The reviewer will be asked to match each test item with the content he/she thinks the item is designed to measure. The assumption is made that if a test has content validity, it will be obvious which items measure specific parts of the content listing. The proportion of items correctly matched to the content listing will be determined for each test (item specification-produced and item bank-produced). A t-test for the significance of the difference between proportions will be performed to see if one type of test has greater content validity than the other type.

The perceived quality of the tests will be determined by sending the two tests in each content area to volunteer upper elementary teachers and curriculum experts in several school districts in Colorado, Michigan and Oregon along with a brief questionnaire (see Attachment B). Reviewers will be asked to respond by saying that Test A is better on the given dimensions, Test B is better, or there is no difference between the two tests. Chi square tests will be computed on the results to see if either test is perceived as significantly better than the other on any of the dimensions.

Description of Instruments

See Attachments A and B for copies of the instruments. The questionnaires were designed by the study's principal investigator. The questionnaire identified as Attachment A should take respondents no more than 30 minutes to complete; the questionnaire identified as Attachment B should take 10-15 minutes to complete. Each district participating in the study can use their own procedures for identifying volunteer upper elementary (grade 3-6) teachers and appropriate curriculum experts. Twenty-five copies of each questionnaire in each subject area (for a total of 100 questionnaires) will be sent to each district's evaluation office with the request that as many reviewers as possible be solicited. Once the reviewers in each subject area are identified, a random half of the group should be sent Attachment A and the other half should be sent Attachment B.

Author and Purpose of Final Report

Authors of the final report are Evelyn Brzezinski (of Interwest Applied Research) and Randy Demaline (of the Northwest Regional Educational Laboratory), both in Portland, Oregon. The final report will be submitted to the National Institute of Education in fulfillment of Grant No. NIE-G-81-0046.

Cost to the District

Besides time costs, the only costs associated with the collection of questionnaire data will be duplication of the instruments and postage to and from the district. All these costs will be covered by the Northwest Regional Educational Laboratory. Sufficient copies of the questionnaires and postage paid return envelopes will be sent to the district's evaluation office; respondents can mail questionnaires directly back to the study investigator in the envelope provided.

Disruption to the Regular Classroom Program

There should be no disruption of the regular class routine since teachers are asked to respond to the questionnaire in their spare time.

Assurance of Anonymity

The district evaluation office will maintain a list of the persons volunteering to participate in the study, but that list will not be provided to the study investigator. Only demographic information (e.g., grade taught, years experience teaching fourth grade) will be requested on the questionnaire.

Timeline for District Involvement

Questionnaires will be sent to the district evaluation office by October 26, 1981. Respondents should return completed questionnaires to the study investigator by November 13, 1981. A copy of the study's final report will be sent to the district in January 1982.

Institutions Associated with the Study

Funding Agency: National Institute of Education

Grant Recipient: Northwest Regional Educational Laboratory

Potential Participating School Districts: Ann Arbor (MI), Beaverton (OR), Boulder Valley (CO), Central Valley (WA), Grand Rapids (MI), Ingham County (MI), Lansing (MI), Portland (OR)

Northwest Regional Educational Laboratory

Memorandum

Date: October 23, 1981

To: Jill Anderson, Marilyn Averill, Dave Kazen, Wayne Neuburger

From: Evelyn Brzezinski EB

Subject: Questionnaires for NIE Study

Here are the 100 questionnaires I told you I would send -- 25 copies of each of 4 different questionnaires. Please try to distribute them as equally as possible; that is, try to send out as many Questionnaire A: Math instruments as you send out Questionnaire B: Math instruments. Some of you have asked if you may send more than one questionnaire to a single teacher. That's perfectly OK, so long as you send one math questionnaire and one reading questionnaire. Under no circumstances should one teacher receive both Questionnaire A and Questionnaire B for the same subject area.

At your convenience, would you please send me a list of the teachers and their addresses to whom you have distributed the questionnaires? I'll need the number of each questionnaire sent out so that I can determine response rates. Also, I'd like to send a thank-you letter to each teacher when this is all over.

Once again, thanks so much for your help with this project. If you get any questions you can't answer, don't hesitate to call me collect at 503/223-3396. Also, please send the list of teachers to me at the following address:

Interwest Applied Research
200 S.W. Market Street, Suite 850
Portland, Oregon 97201

DIRECTIONS TO REVIEWERS

Thank you for agreeing to serve as a test reviewer in this research project. Your responses to this questionnaire will help us learn about the effectiveness of various methods for developing achievement tests. Our goal, of course, is to be able to produce high quality tests in the most efficient and economical method possible. We appreciate your help in this important effort.

The two mathematics tests you have received were developed by different test development methods. Both of the tests were designed to measure the nine topic areas listed on the blue sheet. We would like to know if you see differences in the two tests. Please review each test carefully, then answer the questions.

Please contact Dave Kazen in your district's evaluation office if you have any questions about these directions.

A business reply envelope is attached for your convenience. You need to return just this sheet. You may keep the two tests or throw them away--whatever you would like to do with them is fine. Please mail your questionnaire as soon as possible, but no later than Monday, November 16. Your cooperation is very much appreciated.

1. Look at the nine topic areas which were to be covered on the tests. Which test do you think does a more complete job of covering this content?

   ___ Test A is better ___ Test B is better ___ I see no difference

2. Do you think one test's format (the way the questions are arranged and the order in which they appear) is better than the other?

   ___ Test A is better ___ Test B is better ___ I see no difference

3. These tests are both designed for end-of-fourth-grade pupils. Do you think one test is more appropriate for these students than the other?

   ___ Test A is better ___ Test B is better ___ I see no difference

   If one test is better, why? ______________________________

4. As an overall judgment, do you think one test is better than the other?

   ___ Test A is better ___ Test B is better ___ I see no difference

   If one test is better, why? ______________________________

   ____

   ____

Please answer the following questions:

1. Your district? ______________________________

2. Grade(s) you teach this year? ______________________

3. Have you ever taught fourth grade? ____ Yes ____ No

MATH TEST CONTENT LISTING

A. Add Whole Numbers Without Regrouping
B. Add Whole Numbers With Regrouping
C. Add Decimals
D. Subtract Whole Numbers Without Regrouping
E. Subtract Whole Numbers With Regrouping
F. Subtract Decimals
G. Multiply Whole Numbers Without Carrying
H. Multiply Whole Numbers With Carrying
I. Divide Whole Numbers Without Remainder

DIRECTIONS TO REVIEWERS

Thank you for agreeing to serve as a test reviewer in this research project. Your responses to this questionnaire will help us learn about the effectiveness of various methods for developing achievement tests. Our goal, of course, is to be able to produce high quality tests in the most efficient and economical method possible. We appreciate your help in this important effort.

On the pages which follow, you will see 90 mathematics test items. On the right side of each sheet, you will see nine topic areas listed. Each of the 90 test items measures 1 of the 9 topic areas. In the space to the left of each item number, print the letter of the topic area you think that item measures. For example:

__A__ 1. 1/2 + 1/4 = _____ a. 1/4 b. 2/4 c. 1/2 d. 3/4

You would write A in the blank, as shown, if A were listed in the topic area as "Addition of Simple Fractions."

If you are not sure which topic area the item measures, put a question mark in the space.

Please contact Dave Kazen in your district's evaluation office if you have any questions about these directions.

A business reply envelope is attached for your convenience so you may return this sheet and the questionnaire directly to the study director. Please mail your questionnaire as soon as possible, but no later than Monday, November 16. Your cooperation is very much appreciated.

Please answer the following questions:

1. Your district? ______________________________

2. Grade(s) you teach this year? ____________________

3. Have you ever taught fourth grade? _____ Yes _____ No

END OF FOURTH GRADE MATH TEST

____ 1.

6,328 - 183 =

A. 6,245
B. 6,211
C. 6,205
D. 6,145

____ 2.

1.3568 - 1.254 =

A. .0028
B. .1028
C. 1.2314
D. 12.314

____ 3.

$$\begin{array}{r} 6{,}043 \\ \underline{-5{,}960} \end{array}$$

A. 83
B. 1,083
C. 183
D. 103

____ 4.

$$\begin{array}{r} 4{,}218 \\ 7{,}564 \\ \underline{+3{,}211} \end{array}$$

A. 13,893
B. 14,883
C. 14,893
D. 14,993

____ 5.

72 ÷ 8 =

A. 64
B. 80
C. 7
D. 9

Topic Areas

A. Add Whole Numbers Without Regrouping
B. Add Whole Numbers With Regrouping
C. Add Decimals
D. Subtract Whole Numbers Without Regrouping
E. Subtract Whole Numbers With Regrouping
F. Subtract Decimals
G. Multiply Whole Numbers Without Carrying
H. Multiply Whole Numbers With Carrying
I. Divide Whole Numbers Without Remainder

END OF FOURTH GRADE MATH TEST

____ 6.

518
x 37

A. 5,180
B. 15,156
C. 555
D. 19,166

____ 7.

40 + 831 + 428 =

A. 1,459
B. 1,279
C. 1,299
D. 1,659

____ 8.

$ 8.37
- .65

A. $7.72
B. $7.82
C. $8.92
D. $8.72

____ 9.

50 x 40 =

A. 90
B. 200
C. 2,000
D. 2,090

____ 10.

15
x78

A. 630
B. 930
C. 1,170
D. 1,270

Topic Areas

A. Add Whole Numbers Without Regrouping
B. Add Whole Numbers With Regrouping
C. Add Decimals
D. Subtract Whole Numbers Without Regrouping
E. Subtract Whole Numbers With Regrouping
F. Subtract Decimals
G. Multiply Whole Numbers Without Carrying
H. Multiply Whole Numbers With Carrying
I. Divide Whole Numbers Without Remainder

END OF FOURTH GRADE MATH TEST

_____ 11.

$$\begin{array}{r}\$\ 60.00\\ +\underline{63.19}\end{array}$$

A. $ 12.319
B. $123.09
C. $123.19
D. $120.19

_____ 12.

$4\overline{)68}$ =

A. 12
B. 27
C. 24
D. 17

_____ 13.

$$\begin{array}{r}38\\ \times\ \underline{4}\end{array}$$

A. 102
B. 152
C. 42
D. 132

_____ 14.

500 - 274 =

A. 235
B. 326
C. 226
D. 374

_____ 15.

$$\begin{array}{r}23\\ \times\underline{30}\end{array}$$

A. 690
B. 69
C. 53
D. 713

Topic Areas

A. Add Whole Numbers Without Regrouping
B. Add Whole Numbers With Regrouping
C. Add Decimals
D. Subtract Whole Numbers Without Regrouping
E. Subtract Whole Numbers With Regrouping
F. Subtract Decimals
G. Multiply Whole Numbers Without Carrying
H. Multiply Whole Numbers With Carrying
I. Divide Whole Numbers Without Remainder

END OF FOURTH GRADE MATH TEST

_____ 16.

57
x42

A. 2,394
B. 342
C. 2,014
D. 99

_____ 17.

$20\overline{)180}$

A. 9
B. 90
C. 81
D. 15

_____ 18.

65
x 4

A. 240
B. 249
C. 260
D. 280

_____ 19.

793
+ 45

A. 738
B. 748
C. 848
D. 838

_____ 20.

8 x 7 =

A. 1
B. 15
C. 56
D. 57

Topic Areas

A. Add Whole Numbers Without Regrouping
B. Add Whole Numbers With Regrouping
C. Add Decimals
D. Subtract Whole Numbers Without Regrouping
E. Subtract Whole Numbers With Regrouping
F. Subtract Decimals
G. Multiply Whole Numbers Without Carrying
H. Multiply Whole Numbers With Carrying
I. Divide Whole Numbers Without Remainder

END OF FOURTH GRADE MATH TEST

_____ 21.

.0554 + .7005 =

A. .7559
B. 1.7559
C. 7.559
D. 7559

_____ 22.

$ 55.00
- 3.95

A. $51.05
B. $51.15
C. $52.05
D. $52.15

_____ 23.

525
-416

A. 109
B. 111
C. 131
D. 141

_____ 24.

2 x 3 x 5 =

A. 10
B. 25
C. 30
D. 45

_____ 25.

82
-61

A. 23
B. 141
C. 143
D. 21

Topic Areas

A. Add Whole Numbers Without Regrouping
B. Add Whole Numbers With Regrouping
C. Add Decimals
D. Subtract Whole Numbers Without Regrouping
E. Subtract Whole Numbers With Regrouping
F. Subtract Decimals
G. Multiply Whole Numbers Without Carrying
H. Multiply Whole Numbers With Carrying
I. Divide Whole Numbers Without Remainder

END OF FOURTH GRADE MATH TEST

_____ 26.

500 - 274 =

A. 236
B. 326
C. 226
D. 374

_____ 27.

409
x 7

A. 343
B. 2,803
C 2,863
D. 416

_____ 28.

32
x 4

A. 76
B. 126
C. 128
D. 138

_____ 29.

21.8
+16.9

A. 3.87
B. 37.17
C. 38.7
D. 37.7

_____ 30.

94
-21

A. 63
B. 73
C. 75
D. 115

Topic Areas

A. Add Whole Numbers Without Regrouping
B. Add Whole Numbers With Regrouping
C. Add Decimals
D. Subtract Whole Numbers Without Regrouping
E. Subtract Whole Numbers With Regrouping
F. Subtract Decimals
G. Multiply Whole Numbers Without Carrying
H. Multiply Whole Numbers With Carrying
I. Divide Whole Numbers Without Remainder

END OF FOURTH GRADE MATH TEST

_____ 31.

```
 518
x 37
```

A. 5,180
B. 15,156
C. 555
D. 19,166

_____ 32.

```
 3.6
- .3
```

A. .3
B. 33
C. .33
D. 3.3

_____ 33.

90 ÷ 18 =

A. 5
B. 12
C. 6
D. 20

_____ 34.

40 + 831 + 428 =

A. 1,459
B. 1,279
C. 1,299
D. 1,659

_____ 35.

```
 395
+  4
```

A. 435
B. 399
C. 391
D. 398

Topic Areas

A. Add Whole Numbers Without Regrouping
B. Add Whole Numbers With Regrouping
C. Add Decimals
D. Subtract Whole Numbers Without Regrouping
E. Subtract Whole Numbers With Regrouping
F. Subtract Decimals
C. Multiply Whole Numbers Without Carrying
H. Multiply Whole Numbers With Carrying
I. Divide Whole Numbers Without Remainder

_____ 36.

$39.95 - $4.92 =

A. $35.83
B. $35.03
C. $34.03
D. $34.93

_____ 37.

$4\overline{)68}$

A. 12
B. 27
C. 24
D. 17

_____ 38.

21.8
+16.9

A. 3.87
B. 37.17
C. 38.7
D. 37.7

_____ 39.

769 - 48 =

A. 817
B. 801
C. 737
D. 721

_____ 40.

864
-202

A. 662
B. 1,066
C. 666
D. 1,062

Topic Areas

A. Add Whole Numbers Without Regrouping
B. Add Whole Numbers With Regrouping
C. Add Decimals
D. Subtract Whole Numbers Without Regrouping
E. Subtract Whole Numbers With Regrouping
F. Subtract Decimals
G. Multiply Whole Numbers Without Carrying
H. Multiply Whole Numbers With Carrying
I. Divide Whole Numbers Without Remainder

_____ 41.

$24\overline{)864}$

A. 18
B. 36
C. 41
D. 32

_____ 42.

$$\begin{array}{r} 276 \\ \times\ \underline{11} \end{array}$$

A. 27,876
B. 522
C. 3,036
D. 287

_____ 43.

$$\begin{array}{r} 409 \\ \times\ \underline{\ \ 7} \end{array}$$

A. 343
B. 2,803
C. 2,863
D. 416

_____ 44.

$$\begin{array}{r} 7,418 \\ 361 \\ +\underline{2,100} \end{array}$$

A. 5,879
B. 9,879
C. 9,889
D. 9,979

_____ 45.

$$\begin{array}{r} 6,549 \\ +\underline{4,815} \end{array}$$

A. 11,364
B. 10,354
C. 13,394
D. 11,354

Topic Areas

A. Add Whole Numbers Without Regrouping
B. Add Whole Numbers With Regrouping
C. Add Decimals
D. Subtract Whole Numbers Without Regrouping
E. Subtract Whole Numbers With Regrouping
F. Subtract Decimals
G. Multiply Whole Numbers Without Carrying
H. Multiply Whole Numbers With Carrying
I. Divide Whole Numbers Without Remainder

END OF FOURTH GRADE MATH TEST

____ 46.

  34.48
-33.923

A. 0.125
B. 0.563
C. 0.443
D. 0.557

____ 47.

$ .66
+.71

A. $13.70
B. $ .37
C. $1.37
D. $137.00

____ 48.

  5.2283
+81.261

A. 6.0409
B. 6.0509
C. 86.4893
D. 86.5893

____ 49.

 41
-25

A. 15
B. 16
C. 25
D. 26

____ 50.

$ 8.37
- .65

A. $7.72
B. $7.82
C. $8.92
D. $8.72

Topic Areas

A. Add Whole Numbers Without Regrouping
B. Add Whole Numbers With Regrouping
C. Add Decimals
D. Subtract Whole Numbers Without Regrouping
E. Subtract Whole Numbers With Regrouping
F. Subtract Decimals
G. Multiply Whole Numbers Without Carrying
H. Multiply Whole Numbers With Carrying
I. Divide Whole Numbers Without Remainder

END OF FOURTH GRADE MATH TEST

_____ 51.

7 x 0 =

A. 0
B. 2
C. 7
D. 14

_____ 52.

$$\begin{array}{r} 38 \\ \underline{\times\ 4} \end{array}$$

A. 102
B. 152
C. 143
D. 132

_____ 53.

$$\begin{array}{r} 52 \\ \underline{-19} \end{array}$$

A. 71
B. 33
C. 43
D. 34

_____ 54.

91 + 367 +820+211 =

A. 2,119
B. 1,419
C. 1,489
D. 1,389

_____ 55.

$$\begin{array}{r} 425 \\ \underline{-247} \end{array}$$

A. 288
B. 178
C. 188
D. 278

Topic Areas

A. Add Whole Numbers Without Regrouping
B. Add Whole Numbers With Regrouping
C. Add Decimals
D. Subtract Whole Numbers Without Regrouping
E. Subtract Whole Numbers With Regrouping
F. Subtract Decimals
G. Multiply Whole Numbers Without Carrying
H. Multiply Whole Numbers With Carrying
I. Divide Whole Numbers Without Remainder

END OF FOURTH GRADE MATH TEST

_____ 56.

863
+103

A. 960
B. 966
C. 760
D. 766

_____ 57.

3.6
- .3

A. .3
B. 33
C. .33
D. 3.3

_____ 58.

62 x 4,000 =

A. 248,000
B. 242,000
C. 32,000
D. 28,000

_____ 59.

12
+13

A. 25
B. 26
C. 35
D. 36

_____ 60.

.2 + 3.7 =

A. .39
B. 5.7
C. 39.
D. 3.9

Topic Areas

A. Add Whole Numbers Without Regrouping
B. Add Whole Numbers With Regrouping
C. Add Decimals
D. Subtract Whole Numbers Without Regrouping
E. Subtract Whole Numbers With Regrouping
F. Subtract Decimals
G. Multiply Whole Numbers Without Carrying
H. Multiply Whole Numbers With Carrying
I. Divide Whole Numbers Without Remainder

END OF FOURTH GRADE MATH TEST

_____ 61.

769 - 48 =

A. 817
B. 801
C. 737
D. 721

_____ 62.

$$\begin{array}{r} 276 \\ \times\ \underline{11} \end{array}$$

A. 27,876
B. 522
C. 3,036
D. 287

_____ 63.

$$\begin{array}{r} 9{,}841 \\ -\underline{2{,}888} \end{array}$$

A. 6,953
B. 6,963
C. 7,053
D. 7,953

_____ 64.

$$\begin{array}{r} 8 \\ \times\underline{7} \end{array}$$

A. 42
B. 56
C. 49
D. 15

_____ 65.

90 ÷ 18 =

A. 5
B. 12
C. 6
D. 20

Topic Areas

A. Add Whole Numbers Without Regrouping
B. Add Whole Numbers With Regrouping
C. Add Decimals
D. Subtract Whole Numbers Without Regrouping
E. Subtract Whole Numbers With Regrouping
F. Subtract Decimals
G. Multiply Whole Numbers Without Carrying
H. Multiply Whole Numbers With Carrying
I. Divide Whole Numbers Without Remainder

END OF FOURTH GRADE MATH TEST

____ 66.

$$\begin{array}{r} 6{,}328 \\ -\ \underline{183} \end{array}$$

A. 6,245
B. 6,211
C. 6,205
D. 6,145

____ 67.

$$\begin{array}{r} 57 \\ \underline{\times 42} \end{array}$$

A. 2,394
B. 342
C. 2,014
D. 99

____ 68.

$72 \div 8 =$

A. 64
B. 80
C. 7
D. 9

____ 69.

$$\begin{array}{r} 450 \\ 645 \\ \underline{+866} \end{array}$$

A. 1,851
B. 1,961
C. 1,861
D. 2,461

____ 70.

$783 \div 9 =$

A. 84
B. 85
C. 86
D. 87

Topic Areas

A. Add Whole Numbers Without Regrouping
B. Add Whole Numbers With Regrouping
C. Add Decimals
D. Subtract Whole Numbers Without Regrouping
E. Subtract Whole Numbers With Regrouping
F. Subtract Decimals
G. Multiply Whole Numbers Without Carrying
H. Multiply Whole Numbers With Carrying
I. Divide Whole Numbers Without Remainder

END OF FOURTH GRADE MATH TEST

_____ 71.

.8 - .3 =

A. .05
B. .5
C. 5
D. 5.5

_____ 72.

$35\overline{)3,535}$ =

A. 11
B. 101
C. 110
D. 111

_____ 73.

50 x 40 =

A. 90
B. 200
C. 2,000
D. 2,090

_____ 74.

835.06 + 178.1 =

A. 1,013.07
B. 1,013.16
C. 8,428.7
D. 8,528.7

_____ 75.

6 x 7 x 5 =

A. 425
B. 225
C. 210
D. 47

Topic Areas

A. Add Whole Numbers Without Regrouping
B. Add Whole Numbers With Regrouping
C. Add Decimals
D. Subtract Whole Numbers Without Regrouping
E. Subtract Whole Numbers With Regrouping
F. Subtract Decimals
G. Multiply Whole Numbers Without Carrying
H. Multiply Whole Numbers With Carrying
I. Divide Whole Numbers Without Remainder

END OF FOURTH GRADE MATH TEST

_____ 76.

$$\begin{array}{r} 864 \\ -\underline{202} \end{array}$$

A. 662
B. 1,066
C. 666
D. 1,062

_____ 77.

$$\begin{array}{r} 602 \\ \times\underline{3} \end{array}$$

A. 186
B 1,806
C. 605
D. 1,836

_____ 78.

$$\begin{array}{r} 380 \\ -\underline{295} \end{array}$$

A. 85
B. 95
C. 105
D. 115

_____ 79.

6.7 - .5 =

A. .62
B. .2
C. 6.2
D. 62

_____ 80.

$$\begin{array}{r} 834 \\ +\underline{166} \end{array}$$

A. 900
B. 990
C. 1,000
D. 1,100

Topic Areas

A. Add Whole Numbers Without Regrouping
B. Add Whole Numbers With Regrouping
C. Add Decimals
D. Subtract Whole Numbers Without Regrouping
E. Subtract Whole Numbers With Regrouping
F. Subtract Decimals
G. Multiply Whole Numbers Without Carrying
H. Multiply Whole Numbers With Carrying
I. Divide Whole Numbers Without Remainder

END OF FOURTH GRADE MATH TEST

_____ 81.

$$\begin{array}{r} .6 \\ +\underline{\ .3} \end{array}$$

A. .9
B. 9
C. .09
D. .3

_____ 82.

176.3 - 2.15 =

A. 18.48
B. 17.415
C. 154.8
D. 174.15

_____ 83.

$$\begin{array}{r} 846 \\ \times\underline{\ \ 9} \end{array}$$

A. 7,614
B. 855
C. 1,723
D. 7,254

_____ 84.

$$\begin{array}{r} 602 \\ \times\underline{\ \ 3} \end{array}$$

A. 186
B. 1,806
C. 605
D. 1,836

_____ 85.

6.653 - 3.43 =

A. 3.123
B. 3.223
C. 3.31
D. 6.310

Topic Areas

A. Add Whole Numbers Without Regrouping
B. Add Whole Numbers With Regrouping
C. Add Decimals
D. Subtract Whole Numbers Without Regrouping
E. Subtract Whole Numbers With Regrouping
F. Subtract Decimals
G. Multiply Whole Numbers Without Carrying
H. Multiply Whole Numbers With Carrying
I. Divide Whole Numbers Without Remainder

____ 86.

$ 60.00
+63.19

A. $ 12.319
B. $123.09
C. $123.19
D. $120.19

____ 87.

91
367
820
+211

A. 2,119
B. 1,419
C. 1,489
D. 1,389

____ 88.

37
+85

A. 102
B. 112
C. 122
D. 123

____ 89.

4,068
-2,742

A. 2,326
B. 1,326
C. 2,806
D. 1,306

____ 90.

23
x30

A. 690
B. 69
C. 53
D. 713

Topic Areas

A. Add Whole Numbers Without Regrouping
B. Add Whole Numbers With Regrouping
C. Add Decimals
D. Subtract Whole Numbers Without Regrouping
E. Subtract Whole Numbers With Regrouping
F. Subtract Decimals
G. Multiply Whole Numbers Without Carrying
H. Multiply Whole Numbers With Carrying
I. Divide Whole Numbers Without Remainder

DIRECTIONS TO REVIEWERS

Thank you for agreeing to serve as a test reviewer in this research project. Your responses to this questionnaire will help us learn about the effectiveness of various methods for developing achievement tests. Our goal, of course, is to be able to produce high quality tests in the most efficient and economical method possible. We appreciate your help in this important effort.

The two reading tests you have received were developed by different test development methods. Both of the tests were designed to measure the nine topic areas listed on the blue sheet. We would like to know if you see differences in the two tests. Please review each test carefully, then answer the questions.

Please contact Mike Hunter in your district's evaluation office if you have any questions about these directions.

A business reply envelope is attached for your convenience. You need to return just this sheet. You may keep the two tests or throw them away--whatever you would like to do with them is fine. Please mail you questionnaire as soon as possible, but no later than Friday, November 20. Your cooperation is very much appreciated.

1. Look at the nine topic areas which were to be covered on the tests. Which test do you think does a more complete job of covering this content?

   ___ Test A is better ___ Test B is better ___ I see no difference

2. Do you think one test's format (the way the questions are arranged and the order in which they appear) is better than the other?

   ___ Test A is better ___ Test B is better ___ I see no difference

3. These tests are both designed for end-of-fourth-grade pupils. Do you think one test is more appropriate for these students than the other?

   ___ Test A is better ___ Test B is better ___ I see no difference

4. Do you think there is a difference in terms of the quality of the reading passages included in the tests?

   ___ Test A is better ___ Test B is better ___ I see no difference

5. As an overall judgment, do you think one test is better than the other?

___ Test A is better ___ Test B is better ___ I see no difference

If one test is better, why? ______________________________

____

____

Please answer the following questions:

1. Your district? ______________________________

2. Grade(s) you teach this year? ______________________

3. Have you ever taught fourth [illegible]de? _____ Yes _____ No

READING TEST CONTENT LISTING

A. Use Stated Facts to Answer Question

B. Determine Sequence of Events in a Passage

C. Identify Cause and Effect Relationships

D. Determine Object of a Referent

E. Identify Main Ideas

F. Identify Conclusions and Supporting Facts

G. Distinguish Fact, Opinion, Fiction, Non-Fiction

H. Infer Formulations (Conclusions, Generalizations, Unstated Assumptions)

I. Generate Solutions, Predict Outcomes, Apply Ideas to New Situations

DIRECTIONS TO REVIEWERS

Thank you for agreeing to serve as a test reviewer in this research project. Your responses to this questionnaire will help us learn about the effectiveness of various methods for developing achievement tests. Our goal, of course, is to be able to produce high quality tests in the most efficient and economical method possible. We appreciate your help in this important effort.

On the pages which follow, you will see a number of reading test items. On the right side of each sheet, you will see nine topic areas listed. Each of the 90 test items measures 1 of the 9 topic areas. In the space to the left of each item number, print the letter of the topic area you think that item measures. For example:

__A__ 1. 1/2 + 1/4 = ______ a. 1/4 b. 2/4 c. 1/2 d. 3/4

You would write A in the blank, as shown, if A were listed in the topic area as "Addition of Simple Fractions."

Please contact Mike Hunter in your district's evaluation office if you have any questions about these directions.

A business reply envelope is attached for your convenience so you may return this sheet and the questionnaire directly to the study director. Please mail your questionnaire as soon as possible, but no later than Friday, November 20. Your cooperation is very much appreciated.

- -

Please answer the following questions:

1. Your district? ________________________________

2. Grade(s) you teach this year? ______________________

3. Have you ever taught fourth grade? _____Yes _____No

END OF FOURTH GRADE READING TEST

Topic Areas:
- A. Use Stated Facts to Answer Questions
- B. Determine Sequence of Events in a Passage
- C. Identify Cause and Effect Relationships
- D. Determine Object of a Referent
- E. Identify Main Ideas
- F. Identify Conclusions and Supporting Facts
- G. Distinguish Fact, Opinion, Fiction, Non-Fiction
- H. Infer Formulations (Conclusions, Generalizations, Unstated Assumptions)
- I. Generate Solutions, Predict Outcomes, Apply Ideas to New Situations

Blue skies turned gray. Clouds rolled in. The birds hushed. The hair on Chris's arm stood on end as static electricity built up. The wind whipped the lake into towering waves. It hurled the sand into Chris. "Hurry and get in the car! Wrap your towel around you! Hurry!" someone called. Chris ran to the car, but the sand kept stinging. Tears came to Chris's eyes.

____ 1 **What happened before the birds hushed?**

- **A. Clouds rolled in.**
- **B. The wind started to blow.**
- **C. Chris ran to the car.**
- **D. Tears came to Chris's eyes.**

____ 2 **What happened because of the static electricity?**

- **A. Blue skies turned gray.**
- **B. The hair on Chris's arm stood on end.**
- **C. The wind started to blow.**
- **D. Chris ran to the car.**

____ 3 **How did Chris feel at the end of the story?**

- **A. Happy**
- **B. Friendly**
- **C. Curious**
- **D. Scared**

____ 4 **What do you think will happen next?**

- **A. The sky will turn blue.**
- **B. There will be a rainstorm.**
- **C. Birds will start to sing.**
- **D. There will be an earthquake.**

END OF FOURTH GRADE READING TEST

Topic Areas:
A. Use Stated Facts to Answer Questions
B. Determine Sequence of Events in a Passage
C. Identify Cause and Effect Relationships
D. Determine Object of a Referent
E. Identify Main Ideas
F. Identify Conclusions and Supporting Facts
G. Distinguish Fact, Opinion, Fiction, Non-Fiction
H. Infer Formulations (Conclusions, Generalizations, Unstated Assumptions)
I. Generate Solutions, Predict Outcomes, Apply Ideas to New Situations

The common use of airplanes for travel is fairly new. Airplanes were developed in the early 1900's. They were used as fighters in the two world wars. Until 1950, however, very few people used them for travel. Even then, the planes were small and flying was uncomfortable. Now that we have large, comfortable jets, flying is common. However, it would be fun to fly in an old airplane to see how people used to travel.

____ 5 **This is mainly a story about**

A. jet airplanes.
B. airplanes being used for transportation.
C. uses of airplanes in world wars.
D. the design of comfortable airplanes.

____ 6 **In the 1950's, planes were small and uncomfortable. What effect did this have?**

A. Planes were used as fighters.
B. Jets were developed.
C. Few people traveled by plane.
D. Many people traveled by plane, but they were uncomfortable.

____ 7 **The article states that we now have large, comfortable jets. What result does this have?**

A. Airplanes were developed.
B. The author would like to fly in an old airplane.
C. People buy airplanes.
D. Flying is common.

____ 8 **Which statement is an opinion?**

A. Planes were used as fighters.
B. Now we have large jets.
C. Until 1950 very few people used planes for travel.
D. It would be fun to fly in an old airplane.

____ 9 **In the fourth sentence, them is underlined. To what does the word refer?**

A. fighter airplanes
B. pilots
C. airplanes
D. jet airplanes

END OF FOURTH GRADE READING TEST

Topic Areas:
- A. Use Stated Facts to Answer Questions
- B. Determine Sequence of Events in a Passage
- C. Identify Cause and Effect Relationships
- D. Determine Object of a Referent
- E. Identify Main Ideas
- F. Identify Conclusions and Supporting Facts
- G. Distinguish Fact, Opinion, Fiction, Non-Fiction
- H. Infer Formulations (Conclusions, Generalizations, Unstated Assumptions)
- I. Generate Solutions, Predict Outcomes, Apply Ideas to New Situations

Father's Day was approaching. Emily and Tom wanted to surprise their Dad. They did not have any money for a present, though. "Let's make Dad some brownies," said Emily. "He really likes chocolate, and Mom rarely makes dessert."

"Brownies taste good," said Tom. "I like them a lot."

Tom found a cookbook, and they looked up a recipe. Emily gathered the ingredients. Tom got out bowls, pans, and a mixer. Soon the brownies were mixed and in the oven. Tom set the timer. After 10 minutes, Emily said, "Oh, no, I forgot baking soda. That's what makes the brownies rise!"

____ 10 Why did Tom and Ellen think Dad would like brownies?

A. Tom liked brownies.
B. Father's Day was approaching.
C. Dad likes chocolate.
D. Tom found a cookbook.

____ 11 Who found the cookbook?

A. Mom
B. Dad
C. Tom
D. Emily

____ 12 What happened before Emily found the ingredients?

A. Tom found a cookbook.
B. They mixed the brownies.
C. Tom set the timer.
D. Tom got out bowls.

____ 13 Tom and Ellen are

A. friends.
B. parents.
C. cousins.
D. brother and sister.

____ 14 Why did Tom and Emily bake brownies?

A. They were hungry.
B. They were bored.
C. They liked to cook.
D. They wanted to surprise their Dad.

____ 15 The best title for this passage is:

A. A Father's Day Present
B. Tom Likes Brownies
C. How to Make Brownies
D. Mom Makes A Dessert

____ 16 Which statement is an opinion?

A. Father's Day was approaching.
B. Brownies taste good.
C. Tom found a cookbook.
D. Oh, no, I forgot baking soda.

____ 17 The next time they bake brownies, what should Tom and Ellen do differently?

A. Make more brownies.
B. Double check the ingredients.
C. Save money for a present.
D. Have Mom make more desserts.

____ 18 When Tom and Emily take the brownies out of the oven, they will be

A. high and fluffy.
B. flat.
C. gooey.
D. sugar free.

END OF FOURTH GRADE READING TEST

Topic Areas: A. Use Stated Facts to Answer Questions
B. Determine Sequence of Events in a Passage
C. Identify Cause and Effect Relationships
D. Determine Object of a Referent
E. Identify Main Ideas
F. Identify Conclusions and Supporting Facts
G. Distinguish Fact, Opinion, Fiction, Non-Fiction
H. Infer Formulations (Conclusions, Generalizations, Unstated Assumptions)
I. Generate Solutions, Predict Outcomes, Apply Ideas to New Situations

When the first people settled in the Mohawk River Valley the river was clean and pure. They started towns at a number of places and used the river as a transportation route for people and products.

In the late 1800's and early 1900's, factories were started in these towns. People and factories often have a lot of waste materials to get rid of. Finding a place to put them can be a great problem. Towns and factories along the river answered this problem by dumping wastes into the river. As time went on, more and more wastes were dumped into the river. Some of these wastes were chemicals that were harmful to the animal and plant life existing in the river. The river became dirtier and dirtier until the water was no longer fit to drink; few fish could live in it; and it was not safe for people to swim in. Now that people have realized this, they are attempting to stop dirtying the river so that it is once again fit to use.

____ 19 Which of the following sentences best states the main idea of this passage?

A. When the first settlers made their homes along the Mohawk River, it was clean and pure.
B. Our rivers are being kept clean and pure.
C. No one realizes what made the Mohawk River dirty.
D. The Mohawk River has become unfit for plant and animal life and harmful for people to use.

____ 20 According to the passage, the present condition of the Mohawk River is a result of

A. the activities of the Indians who used the river for transportation before the white man came.
B. the fish dying in the river and making it dirty.
C. too many people using the river for transportation.
D. man using the river as a dumping area for wastes.

____ 21 The Mohawk River was a clean river in the 1800's because

A. it was a young river.
B. there were strong anti-pollution laws.
C. there were no people or towns around to pollute the river.
D. chemicals hadn't been invented yet.

____ 22 What does the story indicate will be necessary to stop pollution of the Mohawk River?

A. Anti-pollution laws and efforts must be made.
B. Factories will have to move to other areas, away from the river.
C. The Valley will have to be abandoned.
D. There will have to be a lot of picketing.

____ 23 The author believes

A. rivers would remain clean if factories stayed away.
B. rivers can be kept clean and safe if wastes are not dumped into them.
C. fish cannot live in rivers where towns are nearby.
D. modern factories pollute the river and the air.

END OF FOURTH GRADE READING TEST

Topic Areas:
- A. Use Stated Facts to Answer Questions
- B. Determine Sequence of Events in a Passage
- C. Identify Cause and Effect Relationships
- D. Determine Object of a Referent
- E. Identify Main Ideas
- F. Identify Conclusions and Supporting Facts
- G. Distinguish Fact, Opinion, Fiction, Non-Fiction
- H. Infer Formulations (Conclusions, Generalizations, Unstated Assumptions)
- I. Generate Solutions, Predict Outcomes, Apply Ideas to New Situations

There was a king who liked to sit outside and look at the North Star. One night when the star was very bright, the king said, "I must go there. I will think of a way."

The next day the king sent for the best builder in his kingdom. "I want you to build a tower to the North Star," said the king.

The builder asked for boxes of all sizes. His helpers put one box on top of another until the pile was very high.

Then the king began to climb the tower. He climbed higher and higher until he came to the top. But the top of the tower did not reach the star. The king needed one more box. So the builders took a box from the bottom of the pile.

____ 24 Which of these things happened first?

- A. The king climbed the tower.
- B. The king called his builder.
- C. The king looked at the North Star.
- D. The king needed another box.

____ 25 What probably happened at the end of the story?

- A. The king got to the North Star.
- B. The builder finished the tower.
- C. The boxes were painted.
- D. The tower and the king fell down.

____ 26 In the third paragraph, His is underlined. Who does His refer to?

- A. the builder
- B. the king
- C. both the builder and the king
- D. the builder's helpers

____ 27 In the fourth paragraph, he is underlined. Who does he refer to?

- A. the builder
- B. the king
- C. the builder's helpers
- D. the box maker

____ 28 This is a story about a king who

- A. wanted to touch the sky.
- B. had a dream about reaching the North Star.
- C. wanted to visit the North Star.
- D. reached the North Star.

END OF FOURTH GRADE READING TEST

Topic Areas:
A. Use Stated Facts to Answer Questions
B. Determine Sequence of Events in a Passage
C. Identify Cause and Effect Relationships
D. Determine Object of a Referent
E. Identify Main Ideas
F. Identify Conclusions and Supporting Facts
G. Distinguish Fact, Opinion, Fiction, Non-Fiction
H. Infer Formulations (Conclusions, Generalizations, Unstated Assumptions)
I. Generate Solutions, Predict Outcomes, Apply Ideas to New Situations

Buy Piggaloo!

Your kids will play it for hours and hours. Make your children better spellers. Two to four can play. For children ages 7 to 12.

_____ 29 Which of the following is a fact?

A. Many people have bought Piggaloo.
B. Children play it for hours and hours.
C. Children who play it are better spellers.
D. Two to four people can play

_____ 30 In the first sentence what does "it" mean?

A. Piggaloo
B. Children
C. Hours
D. Kids

_____ 31 What can parents expect when they buy Piggaloo?

A. Teenagers will like Piggaloo.
B. It takes hours and hours to learn to play.
C. Their children will become better spellers.
D. Three children cannot play.

END OF FOURTH GRADE READING TEST

Topic Areas:
A. Use Stated Facts to Answer Questions
B. Determine Sequence of Events in a Passage
C. Identify Cause and Effect Relationships
D. Determine Object of a Referent
E. Identify Main Ideas
F. Identify Conclusions and Supporting Facts
G. Distinguish Fact, Opinion, Fiction, Non-Fiction
H. Infer Formulations (Conclusions, Generalizations, Unstated Assumptions)
I. Generate Solutions, Predict Outcomes, Apply Ideas to New Situations

"Going to the mountains is going home". This was written by John Muir. Muir was a quite man who loved the mountains. Much of his life was spent trying to preserve the wilderness.

Muir was born in Scotland. In 1868 he moved to California where he lived the rest of his life. Muir loved to spend weeks in the California mountains. He would carry only his notebook, fried bread and tea.

Muir felt that people must have wilderness areas. He convinced many people there was a need. His writing about the wilderness resulted in the establishment of three national parks.

Muir died in 1914, but he is still remembered today. The Sierra Club, a group he founded in 1892, is still devoted to saving the wilderness. Many people continue to read the books he wrote. After his death, a trail was developed and named for him, the John Muir Trail.

32 John Muir was born in

____

A. California.
B. the mountains of Ohio.
C. Scotland.
D. the Sierra Nevada Mountains.

33 What caused the establishment of three national parks?

____

A. Muir started the Sierra Club.
B. John Muir Trail was created.
C. Muir moved to California.
D. Muir wrote about the wilderness.

34 How do we know John Muir loved the wilderness?

____

A. He moved to California.
B. He was a quiet man.
C. People still read his books.
D. He spent weeks in the mountains.

35 Which happened last in the story?

____

A. The Sierra Club was formed.
B. Muir moved to California.
C. Muir died.
D. The John Muir Trail was developed.

36 When was the Sierra Club formed?

____

A. 1868
B. 1892
C. 1914
D. 1920

37 The best title for this story is:

____

A. The Life of John Muir
B. Three National Parks
C. Born in Scotland
D. Fried Bread and Tea

END OF FOURTH GRADE READING TEST

Topic Areas: A. Use Stated Facts to Answer Questions
B. Determine Sequence of Events in a Passage
C. Identify Cause and Effect Relationships
D. Determine Object of a Referent
E. Identify Main Ideas
F. Identify Conclusions and Supporting Facts
G. Distinguish Fact, Opinion, Fiction, Non-Fiction
H. Infer Formulations (Conclusions, Generalizations, Unstated Assumptions)
I. Generate Solutions, Predict Outcomes, Apply Ideas to New Situations

Johnny Appleseed trudged barefooted through the frontier with a burlap bag, holding a treasure of more value than gold, over his shoulder. He carried apple seeds, enough to provide all the territory with trees. The apple was not native to this land, and the seeds had to be imported and planted carefully. That was Johnny's mission--to plant thriving apple orchards throughout the Ohio Valley.

Johnny started with a nursery of his own in Pittsburgh. Then he planted seeds wherever they would grow. He worked deeper and deeper into the frontier country to land where no white man had ever been. So when people moved west and claimed their farms, they found thriving apple orchards waiting for them.

____ 38 Choose the best ending for this sentence:

The first appleseeds had to be imported because

A. the apple trees kept dying.
B. nobody lived deep in the frontier country.
C. the apple did not originally come from America.
D. apples used to be very small.

____ 39 Choose the best conclusion for this passage:

A. Johnny Appleseed wore a coonskin cap in cold weather.
B. No apple seeds were planted in the unsettled country.
C. Johnny Appleseed performed a great service for the settlers in the Ohio Valley.
D. White men lived in all the country to which Johnny Appleseed traveled.

____ 40 Which of the following statements is true?

A. Johnny Appleseed was an Indian.
B. Pittsburgh is east of the Ohio Valley.
C. People who moved into the Ohio Valley were farmers from Pittsburgh.
D. Farmers hired Johnny Appleseed to plant apple orchards for them.

____ 41 When did Johnny Appleseed begin planting apple seeds?

A. After he traveled to Europe.
B. As soon as the apple seeds were sent.
C. Once he had apple seeds from his nursery.
D. After an invitation from the frontier pioneers.

END OF FOURTH GRADE READING TEST

Topic Areas:
- A. Use Stated Facts to Answer Questions
- B. Determine Sequence of Events in a Passage
- C. Identify Cause and Effect Relationships
- D. Determine Object of a Referent
- E. Identify Main Ideas
- F. Identify Conclusions and Supporting Facts
- G. Distinguish Fact, Opinion, Fiction, Non-Fiction
- H. Infer Formulations (Conclusions, Generalizations, Unstated Assumptions)
- I. Generate Solutions, Predict Outcomes, Apply Ideas to New Situations

Ellen knew that Jan's birthday was coming soon. Ellen had saved $4.00 to buy her a present. She knew Jan liked puzzles and model airplanes, but she could not find just the right gift. One day Ellen's mom showed her an ad in the paper.

CLOSE-OUT SPECIAL

Books	$2.98
Puzzles	$3.50
Models	$5.00 and up

____ 42 In the second sentence, "her" refers to

A. Jan.
B. Ellen.
C. Ellen's mom.
D. the newspaper ad.

____ 43 What will Ellen probably do for Jan's present?

A. Buy a book
B. Buy a puzzle
C. Buy a model
D. Give Jan $4.00

____ 44 Which of the following sentences best describes the story?

A. Jan has a birthday.
B. Ellen's mom finds an ad.
C. Ellen looks for a birthday present.
D. Ellen saves $4.00

____ 45 What did Ellen's mom show Ellen?

A. A book
B. A puzzle
C. A model airplane
D. An ad

END OF FOURTH GRADE READING TEST

Topic Areas:
- A. Use Stated Facts to Answer Questions
- B. Determine Sequence of Events in a Passage
- C. Identify Cause and Effect Relationships
- D. Determine Object of a Referent
- E. Identify Main Ideas
- F. Identify Conclusions and Supporting Facts
- G. Distinguish Fact, Opinion, Fiction, Non-Fiction
- H. Infer Formulations (Conclusions, Generalizations, Unstated Assumptions)
- I. Generate Solutions, Predict Outcomes, Apply Ideas to New Situations

We usually think of sugar coming from sugar cane, but it is also made from sugar beets. The sugar beet is an important crop. It is grown in some of our Western states. When the beets are fully grown, the farmers dig them out of the ground. Then they load the beets on trucks and take them to a sugar factory.

Do you know how we get sugar from the sugar beet? At the factory, the beets are washed, sliced, and then placed in hot water. The sugar in the beet chips dissolves in the water. Clear sugar water is drained off and placed in ovens. There it cooks into a thick syrup. Then the syrup goes into large pans. The syrup is boiled until sugar begins to form.

Then the syrup is placed in the bowl on a machine that spins very fast. The liquid is removed, and sugar is left on the sides of the bowl. The sugar is then placed on driers. The driers remove any remaining liquid.

_____ 46 What happens first to the beets at the factory?

A. The beets are placed in hot water.
B. The beets are dug out of the ground.
C. The sugar water from the beets is placed in ovens.
D. The beets are washed.

_____ 47 What is the last thing that is done in the factory to make sugar?

A. The syrup is spun in a machine.
B. The beets are loaded on a truck.
C. The remaining liquid is removed.
D. The sugar water is cooked.

_____ 48 Where are sugar beets grown?

A. South America
B. in our Western states
C. in our Eastern states
D. Hawaii

_____ 49 This story is about

A. where the sugar beet is grown.
B. the machinery used to make beets into sugar.
C. how sugar is spun from beet syrup.
D. how sugar beets become sugar.

_____ 50 The author wants the reader to understand

A. the importance of water in the process of making sugar.
B. farmers have to work hard to make sugar from beets.
C. we wouldn't have any sugar if it weren't for sugar beets.
D. trains are important transportation in the sugar beet industry.

END OF FOURTH GRADE READING TEST

Topic Areas:
- A. Use Stated Facts to Answer Questions
- B. Determine Sequence of Events in a Passage
- C. Identify Cause and Effect Relationships
- D. Determine Object of a Referent
- E. Identify Main Ideas
- F. Identify Conclusions and Supporting Facts
- G. Distinguish Fact, Opinion, Fiction, Non-Fiction
- H. Infer Formulations (Conclusions, Generalizations, Unstated Assumptions)
- I. Generate Solutions, Predict Outcomes, Apply Ideas to New Situations

(We realize this passage and the questions are similar to another set in the questionnaire Please review both sets of items, however.)

The common use of airplanes for travel is fairly new. Airplanes were developed in the early 1900's. However, the planes were small and flying was uncomfortable. Lack of radios made navigation difficult and people were uncertain of their safety. Airplanes were used mainly as fighters in the two World Wars.

In the 1950's larger, more comfortable planes with jet engines were developed. Traveling by plane became common. Today flying is as exciting as it is safe.

____ 51 The story states that we now have large comfortable jets. What result does this have?

- A. Airplanes were developed.
- B. The author would like to fly in an old airplane.
- C. People buy airplanes.
- D. Flying is common.

____ 52 Which statement is an opinion?

- A. Planes were used as fighters.
- B. Now we have large jets.
- C. Very few people used planes for travel in the early 1900's.
- D. Today flying is exciting.

____ 53 In the early days of flying, planes were small and uncomfortable. Navigation was difficult. People worried about safety. What effect did this have?

- A. Planes were used as fighters.
- B. Jets were developed.
- C. Few people traveled by plane.
- D. It would be fun to fly in an old plane.

Topic Areas: A. Use Stated Facts to Answer Questions
B. Determine Sequence of Events in a Passage
C. Identify Cause and Effect Relationships
D. Determine Object of a Referent
E. Identify Main Ideas
F. Identify Conclusions and Supporting Facts
G. Distinguish Fact, Opinion, Fiction, Non-Fiction
H. Infer Formulations (Conclusions, Generalizations, Unstated Assumptions)
I. Generate Solutions, Predict Outcomes, Apply Ideas to New Situations

The old man, a Greek, fought the Roman army to a standstill for nearly 3 years—and almost won. The old man was Archimedes of Syracuse, the greatest scientist of the ancient world.

The Roman army knew his reputation well, and he lived up to it fully. Legend says that when curved mirrors were set up on the walls of Syracuse, a Greek city in Sicily, the besieging Roman ships caught fire. It wasn't sorcery; it was Archimedes. When huge claws were extended outward on beams, ships were caught, raised and overturned. It wasn't magic; it was Archimedes.

____ 54 **According to the passage, which of the following would most cause the ships to burn?**

A. **Strong winds caused the boats to capsize.**
B. **Huge claws caused the ships to burn.**
C. **The Roman ships caught fire because of reflecting mirrors.**
D. **Archimedes hypnotized the sailors.**

____ 55 **If Archimedes were alive today, he would probably be a**

A. **movie producer.**
B. **writer.**
C. **aerospace scientist.**
D. **gardener.**

END OF FOURTH GRADE READING TEST

Topic Areas:
A. Use Stated Facts to Answer Questions
B. Determine Sequence of Events in a Passage
C. Identify Cause and Effect Relationships
D. Determine Object of a Referent
E. Identify Main Ideas
F. Identify Conclusions and Supporting Facts
G. Distinguish Fact, Opinion, Fiction, Non-Fiction
H. Infer Formulations (Conclusions, Generalizations, Unstated Assumptions)
I. Generate Solutions, Predict Outcomes, Apply Ideas to New Situations

They landed with with a "Thump!" Izzy looked at Fizzy. Fizzy slithered off her bed to a small hole in the bottom of the ship. She opened the cover over the hole, crawled down the side of the spaceship and promptly disappeared behind a blade of grass.

"Come quickly, Izzy," she called. "I have never seen this color before! It's like blue, but more yellow. The entire plant is all one color. It doesn't have any leaves, just one tall, slender trunk."

Izzy looked down from the spaceship. "I have studied the Earth," he said. "We are in a huge forest of grass."

56 **Look at the following statements. Then indicate the order in which they occurred according to the story.**

1. **Fizzy crawled down the side of the spaceship.**
2. **Izzy looked at Fizzy.**
3. **The spaceship landed.**
4. **Fizzy slithered to a hole in the ship.**

A. 3, 4, 1, 2
B. 2, 1, 4, 3
C. 1, 2, 4, 3
D. 3, 2, 4, 1

57 **Which animal is Fizzy closest to in size?**

A. An elephant
B. A cat
C. A horse
D. An ant

58 **What will Izzy probably do next?**

A. Disappear behind a blade of grass
B. Crawl down the spaceship
C. Study the earth
D. Take off in the spaceship

59 **In the second paragraph what is Fizzy describing?**

A. A tree
B. Grass
C. Izzy
D. The color blue

60 **What part of the story best supports the idea that Izzy and Fizzy are from outer space?**

A. Fizzy slithered off her bed.
B. Izzy and Fizzy looked at one another.
C. Fizzy crawled down the side of the spaceship.
D. She opened the cover over the hole.

61 **Which of the following statements tells you the story is fiction?**

A. They landed with a "Thump!"
B. Izzy looked at Fizzy.
C. The entire plant is all one color.
D. We are in a huge forest of grass.

END OF FOURTH GRADE READING TEST

Topic Areas: A. Use Stated Facts to Answer Questions
B. Determine Sequence of Events in a Passage
C. Identify Cause and Effect Relationships
D. Determine Object of a Referent
E. Identify Main Ideas
F. Identify Conclusions and Supporting Facts
G. Distinguish Fact, Opinion, Fiction, Non-Fiction
H. Infer Formulations (Conclusions, Generalizations, Unstated Assumptions)
I. Generate Solutions, Predict Outcomes, Apply Ideas to New Situations

On the third day of our vacation, B.D. and I met Jerry. He was standing in a creek and vacuuming rocks, gravel, and sand from the bottom. Then he walked up on the bank and showed us how to pan for gold. First he filled a metal pan with some of the materials he had vacuumed from the creek bottom. Then he scooped in some water. "Slosh it around slowly to let any gold settle," he said. "Then slowly spill out the water and rocks."

I carefully followed his instructions. After ten minutes my pan was almost empty. That's when I saw three small flakes of gold. I was so excited I almost dropped the pan. "Let's come here next year for our vacation," I said to B.D.

____ 62 When panning for gold, what order do you follow?

1. Slosh water in the pan slowly.
2. Spill out rocks and water.
3. Vacuum the river bottom.
4. Put material from the river bottom in a metal pan.

A. 3,1,4,2
B. 2,4,3,1
C. 3,4,1,2
D. 4,1,2,3

____ 63 Why did Jerry slosh water around slowly?

A. So he wouldn't get wet
B. To let any gold settle
C. So he wouldn't kill any fish
D. To keep gravel from falling out

____ 64 Which sentence best summarizes this story?

A. It is important to follow instructions.
B. You can make money on your vacation.
C. I learned to pan for gold.
D. I spent my vacation panning for gold.

____ 65 "B.D." refers to

A. the author's dog.
B. the author.
C. Jerry's friend.
D. the author's friend.

____ 66 If you wanted to pan for gold, you would go to

A. the desert.
B. an abandoned gold mine.
C. a river with a muddy bottom.
D. a river with a rocky bottom.

END OF FOURTH GRADE READING TEST

Topic Areas:
- A. Use Stated Facts to Answer Questions
- B. Determine Sequence of Events in a Passage
- C. Identify Cause and Effect Relationships
- D. Determine Object of a Referent
- E. Identify Main Ideas
- F. Identify Conclusions and Supporting Facts
- G. Distinguish Fact, Opinion, Fiction, Non-Fiction
- H. Infer Formulations (Conclusions, Generalizations, Unstated Assumptions)
- I. Generate Solutions, Predict Outcomes, Apply Ideas to New Situations

Read each sentence. If it tells a fact, choose A. If it tells something that is just a matter of how people think about it, an opinion, choose B.

____ 67 There are more women over 70 than men.

A. fact
B. opinion

____ 68 In Summer we usually have warmer weather than in winter.

A. fact
B. opinion

____ 69 There are 26 letters in the alphabet.

A. fact
B. opinion

____ 70 According to the legend, Johnny Appleseed wore a cooking utensil for a hat.

A. fact
B. opinion

END OF FOURTH GRADE READING TEST

Topic Areas:
- A. Use Stated Facts to Answer Questions
- B. Determine Sequence of Events in a Passage
- C. Identify Cause and Effect Relationships
- D. Determine Object of a Referent
- E. Identify Main Ideas
- F. Identify Conclusions and Supporting Facts
- G. Distinguish Fact, Opinion, Fiction, Non-Fiction
- H. Infer Formulations (Conclusions, Generalizations, Unstated Assumptions)
- I. Generate Solutions, Predict Outcomes, Apply Ideas to New Situations

Loca noticed that the water was getting colder and colder. Like all whales, he was protected from the cold by a layer of fat. But the cold water killed the krill, tiny animals that gray whales eat. "It's time to move south," he thought. He slapped his tail against the water five times. "Let's go" he signaled. The other whales followed him.

As he started to swim, Loca sang a song. It went:

"Sun is low,
Water is cold,
Let's go south to warm water."

Twice a year gray whales like Loca migrate over 3,000 miles. They spend the summer near Alaska where there are many krill to eat. In the winter they swim south to Mexico or Hawaii. The warm water there also contains krill.

The sounds that whales make are called songs. Many people enjoy listening to these songs. Different kinds of whales sing different songs. Gray whale songs are interesting to hear.

---

71 This article is about

A. songs.
B. krill.
C. whales.
D. swimming.

72 The water was getting colder. Krill were being killed. What did this make the whales do?

A. Sing
B. Move north
C. Move south
D. Eat krill

73 Which of the following statements is an opinion?

A. Twice each year gray whales migrate.
B. Whales spend the summer near Alaska.
C. People who study whales have heard them sing.
D. Gray whale songs are interesting to hear.

74 Why do the whales migrate?

A. So they can sing
B. To find krill
C. Because the sun is low
D. They become cold

75 Which of the following happened after Loca slapped his tail against the water?

A. The water got colder.
B. The krill died.
C. Loca sang a song.
D. Loca swam north.

76 Which paragraphs of the story contain fiction?

A. 1 and 2
B. 1 and 3
C. 2 and 3
D. 3 and 4

END OF FOURTH GRADE READING TEST

Topic Areas:
- A. Use Stated Facts to Answer Questions
- B. Determine Sequence of Events in a Passage
- C. Identify Cause and Effect Relationships
- D. Determine Object of a Referent
- E. Identify Main Ideas
- F. Identify Conclusions and Supporting Facts
- G. Distinguish Fact, Opinion, Fiction, Non-Fiction
- H. Infer Formulations (Conclusions, Generalizations, Unstated Assumptions)
- I. Generate Solutions, Predict Outcomes, Apply Ideas to New Situations

As the cat climbed over
the button box
He went one foot at a time

The right forepaw
came
down
safely

But a hind foot caught
In the button box

Buttons
Scattered in all directions.

____ 77 Which of the following sentences tells what probably happened after the buttons were scattered?

A. The cat ran away from the box and buttons.
B. The cat picked up the buttons.
C. The cat fell asleep.
D. The cat fell and hurt himself.

DIRECTIONS TO TEACHERS

The students in your class are being asked to take two tests (either math tests or reading tests, or perhaps both subjects) that were developed by different methods. We want to see if the test statistics are any different once the tests are administered, and so we will analyze the results from the testing with that purpose in mind. The tests are already mixed together in the pile of tests accompanying these directions. Simply distribute a test and an answer sheet to each student in the class and let the students begin.

You might want to review the student directions to make sure that everyone knows how to use the separate answer sheet. Pen or pencil may be used on the answer sheet, and students may write on the test copy. You should return the completed answer sheets to Sandi Meacham Wilson in the district office; she will return answer sheets from all the participating classes to the researchers in Portland. You may keep the tests or throw them away, as you prefer.

Students should be given as much time as they need to answer every question. We estimate that most students will complete the test in under 45 minutes. Students may guess if they are not sure of the correct answer.

Thank you very much for your cooperation. If you have any questions about this testing, please contact Sandi Wilson. She will be receiving a copy of the study results and will share them with you if you are interested. Once again, we appreciate your assistance.

******** MATH TEST ********

STUDENT NAME ____________________

TEACHER NAME ____________________

Test Form (circle one) IS IB

1 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
2 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
3 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
4 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
5 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
6 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
7 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
8 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
9 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
10 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
11 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
12 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
13 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
14 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
15 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃

16 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
17 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
18 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
19 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
20 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
21 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
22 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
23 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
24 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
25 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
26 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
27 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
28 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
29 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
30 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃

31 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
32 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
33 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
34 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
35 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
36 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
37 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
38 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
39 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
40 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
41 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
42 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
43 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
44 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
45 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃

****** READING TEST ******

1 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
2 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
3 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
4 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
5 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
6 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
7 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
8 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
9 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
10 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
11 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
12 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
13 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
14 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
15 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃

16 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
17 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
18 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
19 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
20 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
21 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
22 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
23 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
24 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
25 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
26 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
27 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
28 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
29 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
30 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃

31 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
32 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
33 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
34 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
35 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
36 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
37 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
38 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
39 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
40 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
41 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
42 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
43 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
44 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃
45 ⊂A⊃ ⊂B⊃ ⊂C⊃ ⊂D⊃

STUDENT NAME ____________________

TEACHER NAME ____________________

Test Form (circle one) IS IB